INTRODUCTION TO

HIGH ENERGY PHYSICS

Particle Physics for the Beginner

Lee G Pondrom

World Scientific

INTRODUCTION TO
HIGH ENERGY PHYSICS

Particle Physics for the Beginner

Other Related Titles from World Scientific

Structural Aspects of Quantum Field Theory and Noncommutative Geometry
In 2 Volumes
2nd Edition
by Gerhard Grensing
ISBN: 978-981-123-800-0 (Vol. 1)
ISBN: 978-981-123-801-7 (Vol. 2)

New Era for CP Asymmetries: Axions and Rare Decays of Hadrons and Leptons
by Ikaros I Bigi, Giulia Ricciardi and Marco Pallavicini
ISBN: 978-981-3233-07-2

The Soviet Atomic Project: How the Soviet Union Obtained the Atomic Bomb
by Lee G Pondrom
ISBN: 978-981-3235-55-7
ISBN: 978-981-122-137-8 (pbk)

The Adventure of the Large Hadron Collider: From the Big Bang to the Higgs Boson
by Daniel Denegri, Claude Guyot, Andreas Hoecker and Lydia Roos
foreword by Carlo Rubbia, Nobel laureate in Physics 1984
ISBN: 978-981-3236-08-0
ISBN: 978-981-3237-11-7 (pbk)

INTRODUCTION TO
HIGH ENERGY PHYSICS

Particle Physics for the Beginner

Lee G Pondrom
University of Wisconsin–Madison, USA

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Names: Pondrom, Lee G., 1933– author.
Title: Introduction to high energy physics : particle physics for the beginner /
Lee G. Pondrom, University of Wisconsin--Madison, US.
Description: New Jersey : World Scientific, [2022] | Includes bibliographical references and index.
Identifiers: LCCN 2021033762 | ISBN 9789811222092 (hardcover) |
ISBN 9789811223013 (paperback)
Subjects: LCSH: Particles (Nuclear physics)
Classification: LCC QC793.2 .P66 2022 | DDC 539.7/2--dc23
LC record available at https://lccn.loc.gov/2021033762

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Cover image: CDF detector at the Tevatron Collider in its Assembly Hall in preparation for insertion into the p-pbar collision hall. Credit: Fermilab.

Copyright © 2022 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

For any available supplementary material, please visit
https://www.worldscientific.com/worldscibooks/10.1142/11878#t=suppl

Desk Editor: Ng Kah Fee

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

Preface

Elementary Particle Physics has developed over the last 75 years, and is now a mature field. A theoretical framework called the "Standard Model" has been with us about 40 years, and has been confirmed by many experimental tests. There are to be sure some soft spots. Strong binding of light quarks, and the concept of quark mass, is a difficulty that gets some exposure in this book. The neutrino sector is a field of active experimental work, that is not given full treatment here. Confirmation of the expected properties of the newly discovered Higgs boson, discussed in Chapter 8, is ongoing at the CERN Large Hadron Collider, and there may be future storage rings dedicated to this subject.

Particle physics is a very broad subject, and any text is bound to reflect the taste of the author in the choice of material to be covered. This text is devoted to what might be called "First-Order Electroweak Theory". The consequences of unification of weak and electromagnetic forces have been established by many beautiful experiments since its theoretical development in the 1970s. Most of the detailed predictions of the theory are accessible to the beginning student, and it is the goal of this book to inspire the reader to understand and appreciate the subtleties and powers of the first-order theory. Parity violation in the weak interactions leads to many interesting polarization phenomena involving spin 1/2 and spin one particles. Polarization calculations play a major role in this book.

The definitions of the Dirac spinors for electrons and positrons, and the treatment of the spin one $W^{\pm}$ bosons, and how all of this works in a parity violating theory are impressive. In some problems the cover is removed from the machinery, so you can see the wheels going around, and admire the structure that has been created for us.

QCD, the theory of the strong interactions, is not ignored here, but plays only a peripheral role. There are at least two reasons for this. Perturbative QCD is mathematically more difficult than perturbative electroweak theory, and because of the strong coupling, corrections to the first-order calculations are not small, making QCD less accessible. These problems have been largely overcome by Monte Carlo programs like PYTHIA, that accurately model strong interaction phenomena, and are very useful in the interpretation of experimental data from hadron machines. But a several hundred page program in FORTRAN or C++ is not easy to handle in an introductory class. Computer calculations appear occasionally in this text in the problem sets, but you are free to choose what program language you want to use. In this regard I should mention the program ROOT, a free download from CERN, that combines a C++ interpreter with a graphics package, so you can calculate and plot curves in the same environment. ROOT was used to make many of the illustrations in this book.

One of the first choices in a relativistic theory is called the "metric", that is the convention for handling the fact that in four dimensional space–time the length that is Lorentz invariant is $ds^2 = c^2dt^2 - d\vec{x}^{\,2}$, usually handled by introducing a metric tensor, and upper and lower indices for the four vectors, or the opposite sign $ds^2 = d\vec{x}^{\,2} - c^2dt^2$, that is often accommodated by imaginary time idt. Various authors have defended their choice by accusing others of being feeble minded. Emotions can run high when discussing the metric. Sakurai in his well-known text "Advanced Quantum Mechanics" uses imaginary time, and calls anyone who would use a metric tensor in Special Relativity, where you do not really need one, deficient. Martinus Veltman has adopted the same convention as Sakurai, and has dismissed the use of a metric tensor as misguided, if not worse. This book uses the metric tensor, and upper and lower indices on the four vectors. We do have company. This metric was popularized by Bjorken and Drell in *Relativistic Quantum Mechanics*, and is the choice of two texts that are extensively referred to in this book: *Quarks and Leptons* by Halzen and Martin, and *Collider Physics* by Barger and Phillips. Upper and lower indices can be a nuisance, but so can imaginary numbers. There is no easy escape from the minus sign. Our Lorentz transformations are real, and the length of a physical particle four vector is also real: $p^2 = E^2 - \vec{p}^{\,2} = m^2$. Physical observables are of course metric independent, but intermediate formulas can look very different. This means that in order to check the steps in a particular calculation by referring to another source, be sure that all of the conventions are the same.

Our units are called “natural units”, where $\hbar = c = 1$, c being the speed of light, are both dimensionless, so that mass, momentum, and energy all have the same dimensions. Since c is dimensionless, length and time have the same dimensions also, namely $x \sim 1/E$. The uncertainty relations $\Delta x \Delta p_x \sim \hbar$ and $\Delta E \Delta t \sim \hbar$ are useful in determining which dimensions go where. Conversions to usual quantum mechanical cgs units are supplied throughout the book and in the Appendix. $\hbar c = 197.3\,\text{eV-nm}$ is especially useful, as it gives the correspondence between energy squared and cross section: $1\ \text{GeV}^{-2} = 0.389$ millibarns.

Choice of variables is important. Can you imagine attempting to solve the wave equation for hydrogen in Cartesian coordinates? The Mandelstam variables $s = (p_1+p_2)^2, t = (p_1-p_3)^2$, and $u = (p_1-p_4)^2$ for a 2 to 2 scattering with four momenta $p_1+p_2 = p_3+p_4$ are universally used in this text, and lead to formulas that are generally much simpler than they would be in terms of other choices of kinematic variables. Stanley Mandelstam introduced the variables that bear his name in a paper about pion–nucleon dispersion relations, that almost nobody remembers. But the variables have proven to be a great boon in simplicity for a wide variety of kinematic formulas, and texts that do not use them exclusively are more difficult to read. The basic formulas for e^-e^- (Møller) scattering, e^-e^+ (Bhabha) scattering, and e^-e^+ annihilation are related by permutation of the Mandelstam variables, so if you solve one of them, you have solved them all.

This text would not exist without the concept of using a diagram to display a scattering or decay process, and attaching to the diagram specific rules for writing down the quantum mechanical amplitude. We owe this development to Richard Feynman. Julian Schwinger is credited with the remark that Feynman has given field theory to the masses. We masses are grateful for the contribution. The formula for Compton scattering of X-rays by electrons was first derived by Klein and Nishina soon after Dirac's relativistic electron theory was introduced. It took the authors about six months to work it out. It is derived in Chapter 4. The Klein-Nishina formula is one of the more difficult derivations in the chapter, but even so it only takes a few pages, and maybe a day or two. That is one of many examples of the labor saving techniques that have been developed to facilitate calculations in quantum electrodynamics. If the number of steps in a calculation can be reduced, the chance for making a mistake is diminished. The traditional form for the Klein–Nishina formula has been retained, so it is one of the few cross sections in the text not written in the Mandelstam variables.

Chapter 3 covers some field theory formalism. In quantum mechanics the space coordinates $\vec{x}$ are observables that obey commutation relations with the time derivatives, and time is a parameter supplied by a clock. In the relativistic theory space and time are mixed, and the distinction is removed by assigning all coordinates parameter status. The field itself replaces the space coordinates, and the usual canonical commutation rule becomes $[\Phi(\vec{x},t),\dot{\Phi}(\vec{y},t)] = i\delta(\vec{x}-\vec{y})$. We introduce the concept of a Lagrangian density, and functional derivatives for the wave equation. We then use the formalism in Chapter 4 to derive a simple example both with the field theory and the Feynman rules to demonstrate that the two procedures are the same, although Feynman's approach requires fewer manipulations. Having done that, the formalism is not used again until Chapter 8, where we need it to explain the non-Abelian gauge theory and the Higgs mechanism that are necessary for predictions regarding the properties of the Higgs boson. Non-Abelian gauge theory is not easy, but we only take it far enough to obtain important parts of the electroweak theory.

The quark model appears throughout the text. Chapter 1 has an introduction to the SU(3) group, and the original three quarks (u, d, s). The heavy charmed and bottom quarks are also introduced in Chapter 1. The production and decay of the interesting 175 GeV mass top quark are discussed in Chapter 7. The light quarks gave birth to the quark model, but it is the heavy quarks that confirm its reality.

Quark binding and CP violation are covered in Chapter 9. We harken back to the non-relativistic Schrödinger equation to discuss positronium, and use QED to calculate the rate of ${}^1S_0$ positronium decaying into two γ rays. We introduce the relation between the non-relativistic limit of the relativistic matrix element calculated by the Feynman rules and the corresponding potential in the Schrödinger equation. Then we apply similar ideas to charmonium and bottomonium, where the effective potential is not known. Several non-relativistic potentials are discussed using the WKB approximation. Applications of ordinary quantum mechanics to positronium appear in the problems to Chapter 9. CP violation in the $K^0\bar{K}^0$ system is described in terms of the mass mixing matrix, and the $B\bar{B}$ heavy meson system is analyzed in the same way. The heavy quark analysis is based on a complex phase in the Cabibbo–Kobayashi–Maskawa 3×3 unitary quark mixing matrix. The predictions are compared to recent experimental results from asymmetric e^-e^+ colliders.

The last half of this book was written during the COVID-19 pandemic, that was totally disruptive to ordinary life. It is still raging as I write this in December 2020. The pandemic has shut down everything, and this has had effects on the book. A positive effect has been the forced isolation with minimum distractions to work on the manuscript. The negative effects have been the lack of contact with colleagues, and the access only to publications available online. Two colleagues from the University of Wisconsin, Vernon Barger and Richard Prepost, have helped clarify several points in the course of the book. I owe thanks to Cyrena, my wife, who has had to put up with many weeks of absence in a long career in experimental particle physics.

I have already mentioned Sakurai, Halzen and Martin and Barger and Phillips as valuable sources. I should add Otto Nachtmann's *Elementary Particle Physics*, and Peskin and Schroeder's *Quantum Field Theory* to that list, and the reviews published periodically by the Particle Data Group. I spent over 25 years in the CDF Collaboration at the Tevatron $p\bar{p}$ collider at Fermilab, and that experience shows in many places in the text. The cover illustration is an early version of the CDF detector in the Assembly Hall of the collider. There are references throughout the text to the original literature, principally *Physical Review D*, and *Physical Review Letters.*

Any errors in the text are mine alone. Only a small part of this presentation is truly original. With the exception of one or two complicated kinematic integrals I have worked through every formula in the book, and I give references to the original sources.

It is my hope that the interested reader will gain the skills and confidence to appreciate the richness and complexity of our modern electroweak theory.

Contents

Preface v

Chapter 1. Introduction 1

1.1 Early History . . . 1
1.2 Symmetries and Conservation Laws . . . 3
1.3 Charge Independence . . . 4
1.4 Strange Particles . . . 6
1.5 Two Neutrinos . . . 13
1.6 Unitary Symmetry . . . 14
1.7 Color . . . 29
1.8 Applications of Unitary Symmetry . . . 30
1.9 The Neutral K Mesons . . . 30
1.10 Electron–Positron Storage Rings . . . 34
1.11 Luminosity . . . 35
1.12 Natural Units . . . 35
1.13 Invariant Kinematic Variables . . . 37
1.14 Basic e^-e^+ Reactions . . . 38
1.15 e^-e^+ Annihilation and the Quark Model . . . 39
1.16 Quark Masses . . . 44
1.17 Problems . . . 47
References . . . 52

Chapter 2. Special Relativity 53

2.1 Rotations 53
2.2 Lorentz Transformations 56
2.3 Momentum and Energy 60
2.4 Relativistic Kinematics 62
2.5 Relativistic Wave Equations 69
2.6 Dirac's Equation 71
2.7 Dirac's Equation in an Electromagnetic Field 77
2.8 Lorentz Transformation of Dirac's Equation 86
2.9 Gordon Decomposition of the Vector Current 91
2.10 Charge Conjugation and TCP 92
2.11 Problems 93
References 101

Chapter 3. A Little Field Theory 103

3.1 Introduction 103
3.2 Harmonic Oscillator 104
3.3 Lagrangian Density 105
3.4 Quantization of the Field 108
3.5 Classical Electromagnetic Field 116
3.6 Four-Dimensional Form 118
3.7 Quantization of the Electromagnetic Field 121
3.8 Cross-Section Calculations in Schrödinger's Quantum Mechanics 124
3.9 Rutherford Scattering 129
3.10 Green's Function for Helmholtz's Equation 131
3.11 Time Ordering 134
3.12 Quantization of the Dirac Field 138
3.13 Problems 147
References 157

Chapter 4. Quantum Electrodynamics 159

4.1 S and T Matrices 159
4.2 Transition Rates and Cross-Sections 160

4.3 $e^- + e^+ \to \mu^- + \mu^+$ 163
4.4 Trace Theorems 170
4.5 $e^-e^+ \to \mu^-\mu^+$ Cross-Section 172
4.6 $e^-\mu^- \to e^-\mu^-$ 179
4.7 Feynman Diagrams and Feynman Rules 181
4.8 Bhabha Scattering $e^-e^+ \to e^-e^+$ 183
4.9 $e^-e^- \to e^-e^-$, Møller Scattering 186
4.10 Spin Effects 190
4.11 Compton Scattering $\gamma + e^- \to \gamma + e^-$ 196
4.12 Problems 210
References 212

Chapter 5. Electrodynamics with Protons and Neutrons 213

5.1 Elastic Electron Scattering in the Laboratory Frame 213
5.2 Electron–Proton Elastic Scattering 217
5.3 The Rosenbluth Formula 220
5.4 Electron–Proton Inelastic Scattering 224
5.5 The Parton Model 238
5.6 Parton Distribution Functions for Protons and Neutrons 242
5.7 Problems 252
References 255

Chapter 6. Weak Interactions 257

6.1 Nuclear Beta Decays 257
6.2 Muon Decay $\mu^- \to e^- + \nu_\mu + \bar{\nu}_e$ 269
6.3 Polarized Muons 272
6.4 Pion Decay 277
6.5 Weak Decays of Strange Particles 281
6.6 Hadrons with Charm and Bottom Quarks 292
6.7 Problems 302
References 308

Chapter 7. Electroweak Phenomenology 309

7.1 Introduction to Neutrino Scattering 309
7.2 Neutrino Electron Scattering 311
7.3 Neutrino Quark Scattering 315
7.4 $W^{\pm}$ Intermediate Vector Bosons 322
7.5 Quark Mixing in Coupling to $W^{\pm}$ 338
7.6 Top Quark 341
7.7 Neutral Currents and Neutrinos 347
7.8 Z Production and Decay 355
7.9 Z Production in $p\bar{p}$ Collisions 364
7.10 Z Decay Modes 366
7.11 Problems 369
References 378

Chapter 8. Electroweak Theory 381

8.1 Introduction 381
8.2 Non-Abelian Gauge Transformations 383
8.3 The SU(2) × U(1) Lagrangian 388
8.4 The Higgs Mechanism 394
8.5 Higgs Couplings 402
8.6 W Magnetic Moment 415
8.7 Problems 418
References 420

Chapter 9. Heavy Quark Bound States, Mixing, and CP Violation 421

9.1 Hydrogen Atom 421
9.2 Positronium 426
9.3 Other Exotic Coulomb Atoms 433
9.4 Bound States of $Q\bar{Q}$ Pairs 434
9.5 K^0 Mixing 448
9.6 CP Violation for $K^0, \bar{K}^0$ Mesons 454
9.7 Mixing and CP Violation in $\Upsilon(4S) \to B\bar{B}$ 464
9.8 Problems 485
References 496

Appendix 499

A.1 Numerical Constants . . . 499
A.2 Maxwell's Equations . . . 499
A.3 Special Relativity . . . 501
A.4 Relativistic Kinematics . . . 503
A.5 Dirac δ Functions . . . 504
A.6 Phase Space . . . 505
A.7 Cross-Section and Decay Rate Formulas . . . 509
A.8 Dirac Electron Theory . . . 510
A.9 The Feynman Rules . . . 513
A.10 Trace Theorems . . . 520
A.11 Spin Effects . . . 522
References . . . 524

Index 525

This page intentionally left blank

Chapter 1

Introduction

1.1. Early History

Elementary particle physics is over 100 years old, and is a mature field. An elementary particle may be defined as a constituent of matter that is not a composite of other objects. The birth of the field may be traced to the discovery of natural radioactivity by Henri Becquerel in 1896. Soon thereafter in 1897 J. J. Thomson identified the electron as a constituent of atoms, and Ernest Rutherford identified three distinct types of radioactivity that he called α, β, and γ rays as the first three letters of the Greek alphabet. Alpha particles are helium nuclei, a fact established by Rutherford himself, and are not elementary, but β rays are electrons, and they retain a place as an elementary particle today. A second particle to arrive on the scene was the photon, or light quantum, first proposed by Max Planck in 1900 to account for the frequency spectrum of black body radiation, and then adopted by Albert Einstein in 1905 as a particle to explain the photoelectric effect. Rutherford's γ rays are photons, the force field for electromagnetism.

Atoms are not elementary. They are composed of Z negatively charged electrons orbiting a very small nucleus. Almost all of the atomic mass resides in the nucleus. The nucleus carries a positive charge Z, so that the atom is electrically neutral. The nucleus is not elementary either; it is composed of Z protons and N neutrons for a total weight of $A = N + Z$ nucleons. Neutrons carry no electric charge. Some nuclei are unstable, and decay, emitting an electron and leaving a daughter nucleus behind with the same $A = (N - 1) + (Z + 1)$, i.e. a neutron has turned into a proton. Nuclei heavier than lead ($Z = 82$) all decay by α emission, leaving a daughter nucleus with $A = (N - 2) + (Z - 2)$. The alpha particles have fixed energy, determined by the difference in binding energy of the parent and daughter nuclei. In contrast, the electrons from nuclei that decayed have a continuous spectrum of energies, with an upper endpoint determined by the difference in binding energy of parent and daughter. Since no other particle was observed in the final state, there was an apparent violation of energy conservation. Wolfgang Pauli in 1930 resolved the paradox by postulating that a second unseen neutral particle of very small mass was emitted together with the electron. Fermi called the particle the neutrino — little neutral one — with the symbol ν.

Nuclear physicists in the early 1930s who believed Pauli's hypothesis might list five elementary particles: electron, photon, proton, neutron, and neutrino, labeled e, γ, p, n, and ν respectively. Already in 1928, however, Paul Dirac proposed his famous relativistic electron theory [1], that predicted the antiparticle to the electron, now called the positron, and discovered in a cloud chamber exposed to cosmic rays by Carl Anderson in 1932. At least one elementary particle now had its antiparticle, with the same mass and opposite electric charge.

One of the puzzles of nuclear physics at that time was the nature of the nuclear force responsible for binding the Z protons and N neutrons together, and overcoming the coulomb repulsion between the protons that would cause the nucleus to fly apart. The absence of any long range manifestations of this mysterious force implied that the force had to have a very short range, the order of the size of a nucleus, about 10^{-14} m. As early as 1935 Hideki Yukawa in Kyoto, Japan, proposed that the nucleus is bound by a new massive particle, that we now call the π meson [2]. Then in 1936 Carl Anderson and Seth Neddermeyer, still working with cosmic rays at sea level, identified a new particle of about the correct mass to be the Yukawa meson. Elementary particle physics stalled at this point until a group of Italian physicists studying cosmic rays during World War II showed that the Anderson–Neddermeyer particle, now called a muon, did not interact strongly with a nucleus, and therefore could not be the Yukawa meson [3]. At sea level the cosmic rays are predominantly muons, electrons, and photons. Cecil Powell developed the technique of measuring tracks left by charged particles in photographic emulsion with a microscope, and he exposed the emulsions at high altitudes using mountain tops or balloons. These studies led to the clarification of the muon puzzle. The true Yukawa particle candidate, the π meson or pion, is the parent of the muon: $\pi^+ \to \mu^+ + \nu$ [4]. Here an unseen neutrino was assumed to accompany this decay of an elementary particle, the pion. Electromagnetism, based on the very successful classical description of charges and currents, is one of the fundamental forces of nature. The three forces in decreasing strength are the strong force that binds neutrons and protons in the nucleus, electromagnetism, and the weak force, responsible for decays. The fourth force, gravity, is too weak to play a role in particle physics. There are three charge states of the pion: (π^+, π^0, π^-). All three forces affect the charged pions; they interact strongly with nucleons and other pions. Charged pions decay by weak interaction into a muon and a neutrino. The neutral pion decays electromagnetically into two photons: $\pi^0 \to \gamma + \gamma$, with a decay rate about 10^8 times that of its charged cousin $\pi^+ \to \mu^+ + \nu$. There are two muon charge states μ^- and μ^+, analogous to the electron and positron. By convention the negative particles are called leptons, and the positive ones antileptons. The leptons do not share the strong interaction. Neutrinos are also leptons; ν is paired with e^- and μ^-, and the antineutrino is paired with e^+ and μ^+.[a] Neutrinos, being electrically neutral, have only the weak interaction.

1.2. Symmetries and Conservation Laws

Symmetries and associated conservation laws are important in quantum mechanics. Energy and angular momentum are conserved in classical mechanics, but the consequences of the conservation laws are not as rich as they are in quantum mechanics. An example is the quantum theory of angular momentum, with its group theoretical basis, addition theorems, commutation rules, etc. Quantum mechanics teaches us that any observable that commutes with the Hamiltonian is a constant of the motion. All of this formalism was adopted for elementary particles, and many new symmetries have been discovered that compensated for our lack of knowledge of the forces. One of the early symmetries of the strong interaction is called charge symmetry.

---

[a] ν really comes in two flavors, ν_e and ν_μ, but we are getting ahead of our story.

1.3. Charge Independence

The proton and neutron have almost the same mass, the neutron being 1.3 MeV heavier out of 938 MeV (see Table 1.1). The hypothesis of charge independence states that the nuclear force does not depend on the charge of the nucleon, and the proton and neutron are two charge states of a 'nucleon'. This is analogous to a spin 1/2 doublet, and the charge independence scheme is called I-spin. Nucleons have I-spin 1/2, $|I_z, p\rangle = +1/2$, $|I_z, n\rangle = -1/2$.[b] The three pions that share the strong interaction are assigned I-spin 1, with $|I_z, \pi^+\rangle = +1$, $|I_z, \pi^0\rangle = 0$, and $|I_z, \pi^-\rangle = -1$. The charged pions have the same mass, while the pi zero is about 5 MeV lighter. These small mass differences are ascribed to electromagnetic corrections to charge independence. A charge rule was introduced: $Q = I_z + N/2$, where N is the nucleon number, which is one for (p, n), and zero for the pions.[c] Strong interactions can depend on I, but not I_z, while electromagnetic and weak interactions both violate conservation of I. Charge and nucleon number are conserved in strong, electromagnetic, and weak interactions.

The Pauli exclusion principle can be extended to include I-spin. A state with two protons or two neutrons has two identical fermions which must have an antisymmetric wave function under exchange, or quantum numbers $L_J^{2s+1} = S_0^1$, $P_{0,1,2}^3$, D_2^1, $F_{2,3,4}^3$, etc. in standard spectroscopic notation. The I-spin wave function for two p or two n is even, being I-spin 1. So it costs us nothing to multiply all of the wave functions together and require it to be odd under the exchange:

---

[b] This I_z assignment is reversed in nuclear physics so that neutron rich nuclei have positive I_z.

[c] Earlier N was used for the number of neutrons. Here it is the number of nucleons.

$$\psi(p_1, p_2) = \psi_{\text{space}} \times \psi_{\text{spin}} \times \psi_{I\text{-spin}} = -\psi(p_2, p_1). \tag{1.1}$$

For the (n, p) system, however, we get something new. We require $\psi(n, p) = -\psi(p, n)$, with two possible I-spin states: I-spin 1 like (p, p) or (n, n), and I-spin 0, which may be different. The I-spin 0 wave function is odd under exchange, so space times spin is even, or $S_1^3, P_1^1, D_{1,2,3}^3, F_3^1$, etc. — all of the states forbidden to (p, p) or (n, n), The deuteron (n, p) is S_1^3, with a small admixture of D_1^3, and hence has I-spin 0. The dineutron (n, n) does not bind, and the conclusion is that the nuclear force must be spin dependent, and spin 1 is allowed for (n, p) but not for (n, n).

The pion–nucleon system can have $I = 3/2$ or $I = 1/2$, using the Clebsch–Gordan coefficients of angular momentum (see Table 1.2). These coefficients form a real unitary transformation, so that the coefficient connecting $|I, I_z\rangle$ to $|I_1, I_2, I_{1z}, I_{2z}\rangle$ is the same as its inverse. In matrix language $S_{ij} = S_{ji}^{-1}$. Charge independence means that the strong force can depend on total I, but not on I_z, and this has been confirmed by experiment. Pion proton scattering has three distinct reactions at energies below threshold for production of an extra pion, described in the I-spin model as only two complex amplitudes, $A_{3/2}$ and $A_{1/2}$.[d] Elastic scattering

$$\sigma(\pi^+ + p \to \pi^+ + p) = |A_{3/2}|^2, \tag{1.2}$$

$$\sigma(\pi^- + p \to \pi^- + p) = 1/9|A_{3/2} + 2 \times A_{1/2}|^2. \tag{1.3}$$

Charge exchange

$$\sigma(\pi^- + p \to \pi^0 + n) = 2/9|A_{3/2} - A_{1/2}|^2. \tag{1.4}$$

---

[d] The cross-section has a kinematic factor that we have omitted: $\sigma = \pi/k^2 \times |A|^2$. We assume that, at the same center of mass energy, this factor is the same for all three reactions.

This constraint leads to experimental checks on the model. For instance, if $|A_{3/2}| \gg |A_{1/2}|$, then the cross-sections are in the ratio $(\pi^+ + p)$:$(\pi^- + p)$:$(\pi^0 + n) = 9$:1:2, and the total cross-section ratio is 3:1. The derivation of these results is left as an exercise.

WWII weapons R&D, principally the Manhattan Project to build the atomic bomb, and the Radiation Laboratory at MIT to develop radar, led to a post war explosion in government support of basic research in physics, chemistry, and the life sciences. Ernest Lawrence's cyclotrons began to blossom everywhere, and it was soon learned that the higher energy cyclotrons could produce pions in the laboratory for study. This discovery led many physicists to abandon cosmic rays for elementary particle study, and to begin to use man-made accelerators. Cosmic ray work then turned to the composition, energy distribution, and possible cosmic origin of the rays that bombard the earth's upper atmosphere. This study is still ongoing.

High energy protons from charged particle accelerators were used to produce beams of positive and negative pions to study the new phenomena of production and interaction of these Yukawa particles. Pion production led to copious sources of daughter muons that were also studied to probe electromagnetic and weak interactions.

1.4. Strange Particles

Cosmic rays were not quite done yet as a source of new physics. Close on the heels of the pion discovery came the triggered cloud chamber cosmic ray work of Rochester and Butler that gave examples of production and decay of strange particles. Strange particles were produced by strong interaction and decayed by weak interaction, as were the pions, but the new objects decayed into pions [5]. We now write $K_s^0 \to \pi^+ + \pi^-$. Why did this not proceed by strong interaction, with a decay rate 10^{12} times faster? To add to the frustration, the new particles were too massive to be produced by the existing cyclotrons.

A cyclotron has a constant magnetic field that contains the spiral orbit of the accelerating protons until they reach the outer radius. The magnetic field strength and the outer radius of the orbit determine the final proton momentum. The cyclotron at the University of Chicago, one of the largest, weighed 2000 tons and accelerated protons to 1000 MeV/c momentum. A threshold reaction for production of the new strange particles, $\pi^- + p \to \Lambda + K^0$, requires a pion momentum of about 1000 MeV/c. To produce this pion beam a proton momentum of about three times that much is required. Since the total weight of the cyclotron magnet scales roughly as the cube of the final momentum, construction of such a machine was impractical. An alternate design was proposed, in which the proton beam stays at constant radius in a ring of magnets, and the magnetic field is increased (ramped up) synchronously with the accelerated beam. These machines are called synchrotrons, and several have been built of increasing size over the last 50 years.

The reaction $\pi^- + p \to \Lambda + K^0$ has two strange particles in the final state, the Λ hyperon and the K^0 meson. The pair is produced by strong interaction — called associated production — and each decays by weak interaction into hadrons, pions and protons. Thus, $K^0 \to \pi^+ + \pi^-$, and $\Lambda \to \pi^- + p$. The Λ decay is a vivid example of the 'strange particle' puzzle. The final state $(\pi^- + p)$ cannot be reversed to make a Λ by itself. The strong force reaction $\pi^- + p \to \pi^0 + \Lambda$ does not exist. The neutral pion has to be replaced by a K^0. To explain this a new quantum number was introduced, conserved by the strong and electromagnetic interactions, but not by the weak interactions, called 'strangeness'. The I-spin assignments for the strange particles are reversed. The K mesons have $I = 1/2$, and the heavy particles carrying nucleon number, called baryons, have $I = 0$ or $I = 1$. This is taken care of by adding strangeness to the charge rule: $Q = I_z + (N + S)/2$, which works with the strangeness assignments $(K^+ I_z = +1/2, K^0 I_z = -1/2) S = +1$; $(\bar{K}^0 I_z = +1/2, K^- I_z = -1/2) S = -1$, and $(\Lambda, \Sigma^+, \Sigma^0, \Sigma^-) S = -1$. A consequence of this assignment is that K^- is difficult to produce relative to K^+; while $\pi^- + p \to K^+ + \Sigma^-$ is allowed, the reaction $\pi^- + p \to K^- + \Sigma^+$ is forbidden; and production of a pair of K's: $\pi^- + p \to K^+ + K^- + n$ has a higher energy threshold. On the other hand, a beam of K^- mesons has a larger interaction cross-section with matter than does K^+, because K^- can produce hyperons: $K^- + p \to \Lambda + \pi^0$, while similar reactions with K^+ are forbidden.

The state of knowledge of the elementary particles in the early 1960s is summarized in Table 1.1.

The last column in Table 1.1 changes from lifetime to full width depending on the method of measurement. The two are related by Planck's constant: $\Delta E = \hbar/\tau$, where $\hbar = 6.58 \times 10^{-22}$ MeV-s. A lifetime of 10^{-10} s has $c\tau = 3$ cm, so a fast moving particle travels a finite distance before decaying, and the lifetime can be measured from the distribution in decay lengths. The $\pi^0 c\tau \approx 30$ nm, hard to measure even with help from relativity (see Chapter 2); on the other hand, converting 10^{-16} s into a line width gives about 6.6 eV, a very small fraction of the 135 MeV mass. Electromagnetic lifetimes are difficult to measure: too short to time with a clock, and too narrow to measure a line width. Hadrons that decay by strong interaction have lifetimes $\sim 10^{-23}$ s, or line widths $\sim$66 MeV, and hence display a measurable line shape on a mass plot. These particles are called resonances.

Examples of resonances in Table 1.1 are vector mesons, and spin 3/2 baryons that decay by strong interaction. The Σ^0 decays electromagnetically, as do the π^0 and η. The η with a mass of 548 MeV is an interesting particle, and deserves some attention. About 40% of the time, $\eta \to \gamma + \gamma$, just like the π^0. So the η must be even under charge conjugation: $C|\eta\rangle = +|\eta\rangle$. It also has I-spin zero, so G parity (see Problem 1.4) must be even: $G = C \times e^{i\pi I_y}$; $G|\eta\rangle = +|\eta\rangle$. G parity is even for an even number of pions, and is odd for an odd number of pions, and is conserved by the strong interactions. But the $\eta \to \pi^+ + \pi^- + \pi^0$ 23% of the time, and that is an odd number of pions. Decay into two pions is parity forbidden, and could occur only by weak interaction. The conclusion is that the η, although not a strange particle, does not decay by strong interaction. The 3π final state violates G parity, and must be electromagnetic.

The three charged states of the ρ vector meson, mass = 770 MeV, are assigned I-spin one. The ρ is a very broad resonance, with a full width of $\Gamma = 151$ MeV. All charge states of the ρ decay essentially

Table 1.1: Elementary particles around 1960.

Leptons

Name	Symbol	Mass (MeV)	Spin	Interactions	Lifetime (s)
Electron	e	0.511	1/2	em, weak	Stable
Muon	μ	105.6	1/2	em, weak	2.2×10^{-6}
Neutrino	ν	0	1/2	weak	Stable
Antineutrino	$\bar{\nu}$	0	1/2	weak	Stable

Hadrons (strongly interacting particles)

Pseudoscalar Mesons, spin zero

Name	Symbol	Mass (MeV)	Spin/parity	I-spin	Strangeness	Lifetime (s)
Pion	$\pi^{\pm}$	139.6	0^-	1	0	2.60×10^{-8}
	π^0	135.0	0^-	1	0	8.4×10^{-17}
Eta	η	548	0−	0	0	5.5×10^{-19}
Eta′	η'	958	0^-	0	0	0.2 MeV line width
Kaon	K^+	494	0^-	1/2	+1	1.2×10^{-8}
	K^0	498	0^-	1/2	+1	$K_s^0 = 0.89 \times 10^{-10}$
	$\bar{K}^0$	498	0^-	1/2	−1	$K_L^0 = 5.2 \times 10^{-8}$
	K^-	494	0^-	1/2	−1	1.2×10^{-8}

Vector Mesons, spin 1

Name	Symbol	Mass (MeV)	Spin/parity	I-spin	Strangeness	Full width (MeV)
Rho	$\rho^{\pm}, \rho^0$	775.3	1−	1	0	151.
Omega	ω	782.6	1−	0	0	8.4
Phi	ϕ	1020	1−	0	0	4.4
K^*	K^*	892	1−	1/2	+1	50.8
$\bar{K}^*$	$\bar{K}^*$	892	1−	1/2	−1	50.8

Baryons spin 1/2

Name	Symbol	Mass (MeV)	Spin/parity	I-spin	Strangeness	Lifetime (s)
Proton	p	938.3	1/2+	1/2	0	Stable
Neutron	n	939.6	1/2+	1/2	0	879
Lambda	Λ	1115.6	1/2+	0	−1	2.63×10^{-10}
Sigma	Σ^+	1189	1/2+	1	−1	0.8×10^{-10}
	Σ^0	1192.6	1/2+	1	−1	7.0×10^{-20}
	Σ^-	1197.4	1/2+	1	−1	1.5×10^{-10}
Cascade	Ξ^0	1315	1/2+	1/2	−2	2.9×10^{-10}
	Ξ^-	1322	1/2+	1/2	−2	1.6×10^{-10}

Baryons spin 3/2

Name	Symbol	Mass (MeV)	Spin/parity	I-spin	Strangeness	Full width (MeV)
Delta	Δ	1232	3/2+	3/2	0	120
Sigma∗	Σ^*	1385	3/2+	1	−1	36
Cascade∗	Ξ^*	1530	3/2+	1/2	−2	10

100% of the time into two pions: $\rho^0 \to \pi^+ + \pi^-$, similarly for $\rho^\pm$. All three charge states are therefore assigned even G parity. The I-spin singlet ω has very nearly the same mass, $m_\omega = 782\,\mathrm{MeV}$, but has a much narrower width $\Gamma = 8.4\,\mathrm{MeV}$. It decays predominantly (89%) into three pions: $\omega \to \pi^+ + \pi^- + \pi^0$. So it has odd G parity. But 2% of the decays are $\omega \to \pi^+ + \pi^-$. This gives a measure of the validity of G parity, and in turn of I-spin, conservation. The two pion final state of the ω can interfere with the two pion final state of the ρ^0.

The resonance $\Delta^{++} \to p + \pi^+$ with a full width of about 115 MeV, corresponding to a lifetime of about 0.6×10^{-23} s, 10^{-13} times a typical weak interaction lifetime for the strange baryons. The broad line widths mean that the masses are not precisely defined, but the peak mass differences are approximately equal: $m_{\Sigma^*} - m_\Delta = 153\,\mathrm{MeV}$, and $m_{\Xi^*} - m_{\Sigma^*} = 145\,\mathrm{MeV}$. Figure 1.1 shows the total $\pi^+ + p$ and $\pi^- + p$ cross-sections in the resonance energy range. The Δ^{++} spin 3/2, I-spin 3/2 peak and its width are clearly visible. The spin 3/2 assignment comes from a phase shift analysis of the scattering amplitude. The phase shift description of elastic scattering is covered in standard texts on quantum mechanics. The scattering amplitude $f(\theta)$ is expanded in a power series:

$$f(\theta) = \frac{1}{k} \times \sum_l e^{i\delta_l} \sin \delta_l (2l+1) P_l(\cos\theta). \tag{1.5}$$

Here $P_l(x)$ is the polynomial that satisfies Legendre's differential equation:

$$\frac{d}{dx}(1-x^2)\frac{dP_l}{dx} + l(l+1)P_l(x) = 0, -1 \leq x \leq +1. \tag{1.6}$$

The orthogonality of the polynomials for different l can be proved from this equation. Write down the equations for two different l's; multiply the P_l equation by $P_{l'}$, and vice versa, subtract and integrate. All terms vanish except for

$$(l(l+1) - l'(l'+1)) \int_{-1}^{+1} P_l(x) P_{l'}(x) dx = 0. \tag{1.7}$$

Legendre's equation is homogeneous, so any normalization can be adopted. The convention we will use is $P_l(1) = 1$. Then the normalization integral is

$$\int_{-1}^{+1} P_l(x)P_{l'}(x)dx = \frac{2}{(2l+1)} \times \delta_{l,l'}, \tag{1.8}$$

where $\delta_{l,l'} = 1$ if $l = l'$, and zero otherwise. With this choice of normalization the forward scattering amplitude ($\theta = 0$) has the simple form:

$$f(0) = \frac{1}{k} \times \sum_l (2l+1)e^{i\delta_l} \sin\delta_l. \tag{1.9}$$

The parameter δ_l is the phase shift for the lth partial wave. and is real for elastic scattering. The polar angle θ is between the outgoing momentum vector and the incident beam direction, chosen to be the z direction. This formula is for spinless particles. The total cross-section is obtained by integration of $|f(\theta)|^2$ over the solid angle $\mathrm{d}\Omega = \sin\theta\mathrm{d}\theta\mathrm{d}\phi$, and simplifies to a single sum because of the orthogonality of the Legendre polynomials:

$$\sigma = \int \mathrm{d}\Omega|\mathrm{f}(\theta)|^2 = \frac{4\pi}{\mathrm{k}^2}\sum_{\mathrm{l}}(2\mathrm{l}+1)\sin^2\delta_{\mathrm{l}}. \tag{1.10}$$

Note that $(4\pi/k) \times \mathrm{Im}\{f(0)\} = \sigma$. This is called the optical theorem, and relates the forward elastic amplitude to the total cross-section. Although this proof does not include inelastic channels, the theorem holds for $\sigma_{\mathrm{tot}} = \sigma_{\mathrm{elastic}} + \sigma_{\mathrm{inelastic}} = (4\pi/k) \times \mathrm{Im}\{f(0)\}$. The phase shift δ_l becomes complex if there is an inelastic channel, so that $|e^{i\delta_l}| < 1$.

For a resonance with angular momentum $\hbar l$, the cross section is a maximum for that partial wave, $\delta_l = \pi/2$. A partial wave analysis is quite tedious at high energy, where the number of partial waves is large. The short range of the nuclear force gives a cutoff $l_{\mathrm{max}} \sim k \times a$, where k is the wave number, and a is the force range. Particles with higher angular momentum "miss" the target. For example, with $p = \hbar k = 300\,\mathrm{MeV/c}$, and $a = 10^{-15}\,\mathrm{m}$ (the proton radius), $k \times a = 1.5$.

The line shape comes from an analysis of the way the phase shift moves with energy through the resonance $\delta_l = \pi/2$. Dropping the l index, suppose the phase shift is linear in the energy:

$$\delta = \frac{\pi}{2} - \epsilon(E - E_0), \tag{1.11}$$

where E_0 is the resonance energy, and the term $\epsilon(E - E_0)$ is small. Then we have

$$\frac{(e^{i\delta}\sin\delta)}{k} \approx \frac{\epsilon(E - E_0) + i}{k}. \tag{1.12}$$

Now multiply and divide by the complex conjugate $\epsilon(E - E_0) - i$, and drop the quadratic term in the numerator, with $\epsilon = 2/\Gamma$ to obtain the Breit–Wigner form:

$$f(\theta) = (2l + 1) \times \frac{\Gamma}{2k((E - E_0) - i\Gamma/2)} P_l(\cos\theta). \tag{1.13}$$

Note that at the resonance energy the scattering amplitude is pure imaginary. The total cross section for the resonant partial wave through the resonance region is

$$\sigma = \frac{4\pi(2l + 1)}{k^2} \frac{\Gamma^2}{4((E - E_0)^2 + \Gamma^2/4)}. \tag{1.14}$$

The half width at half maximum is $\Gamma/2$. The full width and the lifetime are related by Plank's constant: $\Gamma = \hbar/\tau$. Long lived states are narrow.

Adding Equations (1.3) and (1.4) (the cross terms cancel), gives

$$\sigma(\pi^+ + p) = |A_{3/2}|^2, \tag{1.15}$$

$$\sigma(\pi^- + p) = 2/3|A_{1/2}|^2 + 1/3|A_{3/2}|^2. \tag{1.16}$$

Figure 1.1 shows $\sigma(\pi^+ p) \sim 200$ mb and $\sigma(\pi^- p) \sim 70$ mb at the resonance peak — about 300 MeV/c incident pion beam momentum. Hence the data agree very well with the I-spin hypothesis, and dominance of the $A_{3/2}$ amplitude.

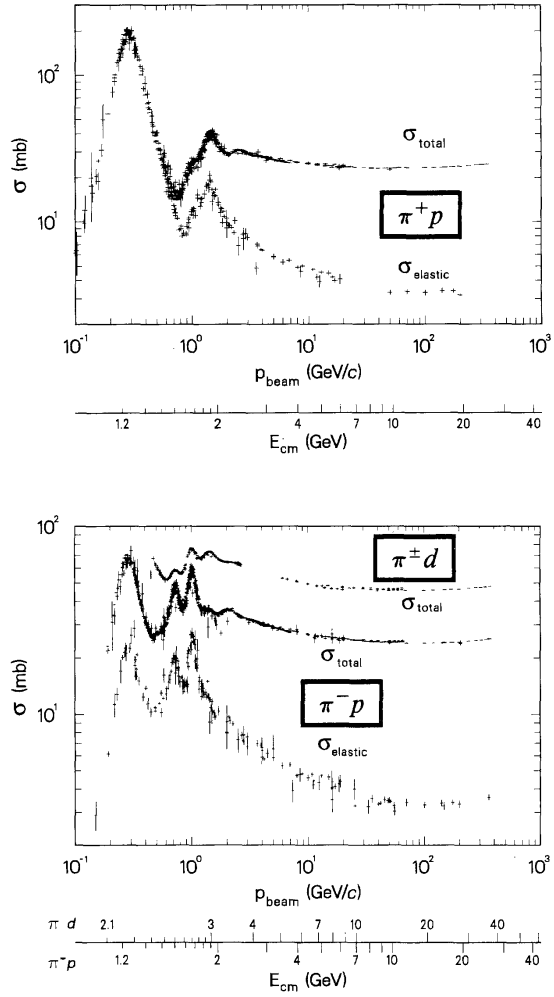


Figure 1.1: Pion–proton total cross-sections as a function of incident pion momentum on a stationary hydrogen target. Energy available in the center of mass is also shown. The strong π^+p resonance is the Δ^{++}.

1.5. Two Neutrinos

There are many good Chinese restaurants in the neighborhood of Columbia University in New York City. At lunch one day about 1960 T. D. Lee, a theoretical physicist at Columbia, asked his colleagues how the weak interaction could be studied at high energy? All of the known weak decays were at low energies. This stimulated a collaboration among Leon Lederman, Mel Schwartz, and Jack Steinberger to use the AGS synchrotron at Brookhaven National Laboratory to produce a beam of high energy neutrinos from pion decay in flight, and to perform the first ever accelerator based neutrino experiment.[e] A proton beam produces more π^+ than π^-, and $\pi^+ \to \mu^+ + \nu$. The neutrinos in turn would produce negative leptons: $\nu + n \to e^- + p$, or $\nu + n \to \mu^- + p$. Aluminum plates supplied the target neutrons, and lots of concrete and steel shielding was required to suppress background radiation from the accelerator. The authors reported 34 single muon tracks that were consistent with neutrino interactions, with a possible background from cosmic rays (measured in separate background runs) of five events, for a signal of 29 muons. If the neutrinos coupled to muons and electrons equally, they would expect 29 electron 'showers', while they observed five shower candidates. A muon in the aluminum plate spark chamber left a straight track, while an electron would create a shower of e^+e^- pairs . The conclusion was that there are two kinds of neutrinos, called ν_μ and ν_e. The muon neutrinos from pion decay produce muons, while the electron neutrinos from nuclear β decay produce electrons. The leptons in Table 1.1 should be modified to indicate two families: $(\mathrm{e}^-, \nu_\mathrm{e})$ and (μ^-, ν_μ), with the corresponding pairs of antileptons.[f]

1.6. Unitary Symmetry

The next step in the study of the strong interacting hadrons was a theoretical arrangement of the particles into multiplets by an extension of I-spin that included strangeness. All particles in one of the new multiplets had to have the same spin, intrinsic parity (discussed below), baryon number, and in the ideal case the same mass, but could have different strangeness and I-spin. The I-spin group, being analogous to quantum mechanical angular momentum, was familiar. The basic unit of the Lie group, called SU(2), is spin 1/2, the doublet. All higher representations can be formed by adding together spin 1/2's. The first step is $1/2 + 1/2$. In terms of multiplicity $2 \otimes 2 = 3 \oplus 1$, where the triplet is I-spin 1, and the singlet is I-spin 0. The group has three operators, I_x, I_y, and I_z. A fourth expression, total I, is formed from the sum $I^2 = I_x^2 + I_y^2 + I_z^2$ with eigenvalues $I(I+1)$. Quantum numbers for I can be integer, including zero, or half integer. The number of states in a representation is the number of values of $I_z = 2I + 1$, odd for integer I, and even for 1/2 integer I. The Clebsch–Gordan coefficients for adding $I_1 = I_2 = 1/2$, $I_1 = 1/2, I_2 = 1$, and $I_1 = 1, I_2 = 1$ are shown in Table 1.2. The phases chosen in the coefficients reflect the definition of the spherical harmonics, and various forms are used. We shall adopt the convention:

---

[e] These gentlemen received the 1988 Nobel Prize in physics for this experiment.

[f] Reactor neutrinos, from $n \to p + e^- + \bar{\nu_e}$, produce positrons, while solar neutrinos, from $p \to n + e^+ + \nu_e$ produce electrons.

$$\begin{aligned}\text{For } m \geq 0,\ Y_l^m(\theta,\phi) &= \sqrt{(2l+1)(l-m)!/(4\pi(l+m)!)}\\ &\quad \times (-1)^m P_l^m(\cos\theta)e^{im\phi}. \qquad (1.17)\end{aligned}$$

Here the associated Legendre function is

$$P_l^m(x) = (1-x^2)^{m/2}\frac{d^m P_l(x)}{dx^m}. \qquad (1.18)$$

For negative m, $Y_l^{-m} = (-1)^m Y_l^{m*}$. Note that this choice of phase makes Y_1^{+1} negative.

The addition of another conserved quantity, strangeness S, or hypercharge $Y = N + S$, requires a larger group, called SU(3). The basic unit of SU(3) is a triplet, 3, and it has a dual, an antitriplet, These states are shown in (I, Y) space in Figure 1.2.

The states at the corners of the 3 are labeled as the quarks: (u, d, s). The quarks have peculiar quantum numbers: (u, d) $I = 1/2$, $I_z = \pm 1/2$ looks OK, but $N = 1/3$, and $S = 0$, so $e_q = +2/3$ for u, and $-1/3$ for d. For the strange quark s, $I = 0$, $S = -1$, $N = 1/3$, giving $e_q = -1/3$. Here e_q is the quark charge in units of the elementary charge e. The reason for assigning $N = 1/3$ to the quarks

Table 1.2: Clebsch–Gordan coefficients for adding $I_1 = 1/2$ to $I_2 = 1/2$; $I_1 = 1/2$ to $I_2 = 1$; and $I_1 = 1$ to $I_2 = 1$.

- $I_1 = 1/2; I_2 = 1/2$

  $I = 1$

  $\langle 1, +1| + 1/2, +1/2\rangle = 1 \qquad \langle 1, -1| - 1/2, -1/2\rangle = 1,$

  $\langle 1, 0| + 1/2, -1/2\rangle = \sqrt{1/2} \qquad \langle 1, 0| - 1/2, +1/2\rangle = \sqrt{1/2}.$

  $I = 0$

  $\langle 0, 0| + 1/2, -1/2\rangle = \sqrt{1/2} \qquad \langle 0, 0| - 1/2, +1/2\rangle = -\sqrt{1/2}.$

- $I_1 = 1/2; I_2 = 1$

  $I = 3/2$

  $\langle 3/2, +3/2| + 1/2, +1\rangle = 1 \qquad \langle 3/2, -3/2| - 1/2, -1\rangle = 1,$

  $\langle 3/2, +1/2| + 1/2, 0\rangle = \sqrt{2/3} \qquad \langle 3/2, +1/2| - 1/2, +1\rangle = \sqrt{1/3},$

  $\langle 3/2, -1/2| - 1/2, 0\rangle = \sqrt{2/3} \qquad \langle 3/2, -1/2| + 1/2, -1\rangle = \sqrt{1/3}.$

  $I = 1/2$

  $\langle 1/2, +1/2| + 1/2, 0\rangle = -\sqrt{1/3} \qquad \langle 1/2, +1/2| - 1/2, +1\rangle = \sqrt{2/3},$

  $\langle 1/2, -1/2| - 1/2, 0\rangle = \sqrt{1/3} \qquad \langle 1/2, -1/2| + 1/2, -1\rangle = -\sqrt{2/3}.$

- $I_1 = 1; I_2 = 1$

  $I = 2$

  $\langle 2, +2| + 1, +1\rangle = 1 \qquad \langle 2, -2| - 1, -1\rangle = 1,$

  $\langle 2, +1| + 1, 0\rangle = \sqrt{1/2} \qquad \langle 2, +1|0, +1\rangle = \sqrt{1/2},$

  $\langle 2, 0|0, 0\rangle = \sqrt{2/3}.$

  $\langle 2, 0| + 1, -1\rangle = \sqrt{1/6} \qquad \langle 2, 0| - 1, +1\rangle = \sqrt{1/6},$

  $\langle 2, -1|0, -1\rangle = \sqrt{1/2} \qquad \langle 2, -1| - 1, 0\rangle = \sqrt{1/2}.$

  $I = 1$

  $\langle 1, +1| + 1, 0\rangle = \sqrt{1/2} \qquad \langle 1, +1|0, +1\rangle = -\sqrt{1/2},$

  $\langle 1, 0| + 1, -1\rangle = \sqrt{1/2} \qquad \langle 1, 0| - 1, +1\rangle = -\sqrt{1/2},$

  $\langle 1, 0|0, 0\rangle = 0,$

  $\langle 1, -1| - 1, 0\rangle = \sqrt{1/2} \qquad \langle 1, -1|0, -1\rangle = -\sqrt{1/2}.$

  $I = 0$

  $\langle 0, 0|1, -1\rangle = \sqrt{1/3} \qquad \langle 0, 0| - 1, 1\rangle = \sqrt{1/3},$

  $\langle 0, 0|0, 0\rangle = -\sqrt{1/3}.$

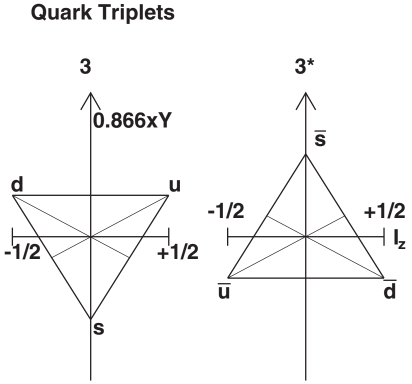


Figure 1.2: Triplet 3 and antitriplet 3* fundamental representations of SU(3) plotted in I–Y space. the vertical axis is $\sqrt{3}/2 \times Y$ to scale the triangles to be equilateral. The corners are labeled in the quarks for the 3 and the antiquarks for the 3*.

will be obvious presently. The antiquarks have all quantum numbers reversed. The (u, d) quarks are the only ones to have non-zero I spin. Several careful experiments, modern versions of the Millikan oil drop experiment for example, have looked for particles with 1/3 integral elementary charge, with no success so far. The quarks have not appeared as free particles in the laboratory. Nevertheless, there is plenty of evidence in support of the quark model, as we shall see.

Addition theorems for the basic triplets are: $3 \otimes 3 = 6 \oplus 3^*$, $3 \otimes 3^* = 8 \oplus 1$, $3 \otimes 6 = 10 \oplus 8$, and $3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1$. The 6 and the 10 are triangular representations, like the 3. There are no known states that have only two quarks, so the 6 is unphysical, but the 10, which is arranged like pool balls $4, 3, 2, 1$, is identified with the spin 3/2 baryon resonances. One of the octets in $3 \otimes 3 \otimes 3$ is identified with the spin 1/2 baryons. This octet and the decuplet form baryons with three quarks, requiring the assignment of N $= 1/3$ to each quark. Figures 1.3 and 1.4 show the baryon octet and decuplet, respectively, while Figures 1.5 and 1.6 show two nonets of mesons from $3 \otimes 3^*$, a pseudoscalar nonet, and a vector nonet. The octet representation allows for a degeneracy in the center, one state with $I_z = Y = 0$,

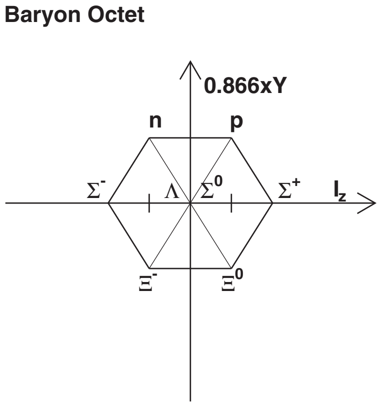


Figure 1.3: Spin 1/2 baryons arranged in the octet representation of SU(3). The six states around the perimeter are non-degenerate. There are two degenerate states in the center: Λ and Σ^0.

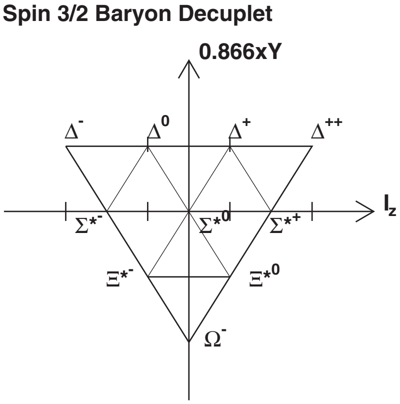


Figure 1.4: Spin 3/2 baryons arranged in the decuplet representation of SU(3). It is a triangular representation, so the states are all non-degenerate. The Ω^- at the bottom tip with $S = -3$ was predicted before its discovery.

and the other with $I = Y = 0$. The ninth particle in the meson family also has $I = Y = 0$. Degeneracy is a feature of the non-triangular representations; the triangles have no degenerate states. States on the boundary of a representation are non-degenerate, and the degeneracy increases by one for each step inside, until a triangle is reached. After that the degeneracy does not change. This scheme is

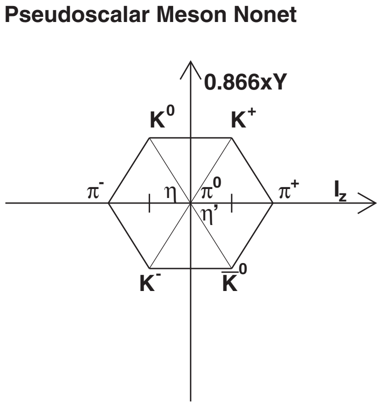

Pseudoscalar Meson Nonet
0.866xY
K0
K+
π-
η
π0
π+
Iz
η'
K-
K̄0

Figure 1.5: Spin-parity 0^- pseudoscalar meson nonet from $3 \otimes 3^* = 8 \oplus 1$. There are three states in the center, one belonging to an *I*-spin triplet, the other two with *I*-spin zero.

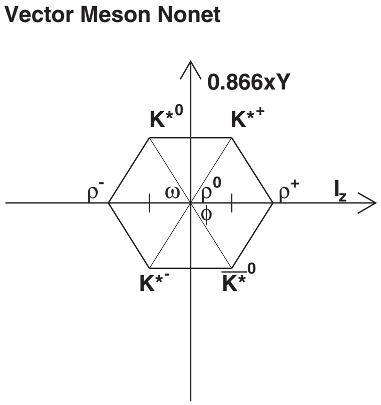

Vector Meson Nonet
0.866xY
K*0
K*+
ρ-
ω
ρ0
ρ+
Iz
φ
K*-
K̄*0

Figure 1.6: Spin-parity 1^- vector meson nonet. The configuration of states is the same as the pseudoscalars.

called unitary symmetry. The number of states in a representation of SU(3) are $3, 6, 8, 10, 15, 27, \ldots$ The 8 and the 27 are symmetrical, and have degenerate states in the middle (the 27 has three), but no duals. The triangular representations have duals that are plotted inverted, as they have all quantum numbers reversed: 3*, 6*, 10*, 15*. There are two different 15's, one prism shaped, so the dimensionality alone is not a unique signature of the representation. Two integers are required to specify a representation uniquely, written (p, q), with the dual representation interchanged (q, p). The dimensionality is given by the formula $\mathrm{N}(p, q) = (1 + p)(1 + q)(1 + (p + q)/2)$, and p and q can be any integer. Symmetrical representations have $p = q$ and no dual. Quarks are $(1, 0)$, anti quarks are $(0, 1)$, and the octet is $(1, 1)$. The two 15's are $(4, 0)$ and $(2, 1)$. Unlike SU(2), there is not a representation for every integer; there is no 5, for example. A specific 3×3 quark-based representation of the eight generators of the group is given in Equation (1.19). Analogous to I^2 in SU(2), there are two combinations of the eight generators that commute with all of them, called Casimir operators. The two integers (p, q) specify these diagonal operators, and uniquely label the representations.

The SU(3) analog of the Pauli spin matrices was introduced by Gell-Mann. There are eight 3×3 matrices, two of which are diagonal. The first three matrices are the Pauli matrices for I-spin 1/2, while the other matrices mix I and Y. They are

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}; \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}; \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix};$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}; \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{1.19}$$

The rows and columns of the matrices refer to the quarks (u, d, s). The first three matrices are the Pauli matrices for the I-spin 1/2 u and d quarks. The commutation rules are satisfied by $\lambda_i/2$, and are characterized by structure constants f_{ijk}:

$$[\lambda_i/2, \lambda_j/2] = i f_{ijk} \lambda_k/2. \tag{1.20}$$

These commutation rules define the algebra, and are satisfied by operators for any representation. The basic 3×3 matrices, like the

Table 1.3: SU(3) structure constants.

(ijk)	f_{ijk}	(ijk)	d_{ijk}	(ijk)	d_{ijk}
123	1	118	$\frac{1}{\sqrt{3}}$	366	$-\frac{1}{2}$
147	$\frac{1}{2}$	146	$\frac{1}{2}$	377	$-\frac{1}{2}$
156	$-\frac{1}{2}$	157	$\frac{1}{2}$	448	$-\frac{1}{2\sqrt{3}}$
246	$\frac{1}{2}$	228	$\frac{1}{\sqrt{3}}$	558	$-\frac{1}{2\sqrt{3}}$
257	$\frac{1}{2}$	247	$-\frac{1}{2}$	668	$-\frac{1}{2\sqrt{3}}$
345	$\frac{1}{2}$	256	$\frac{1}{2}$	778	$-\frac{1}{2\sqrt{3}}$
367	$-\frac{1}{2}$	338	$\frac{1}{\sqrt{3}}$	888	$-\frac{1}{\sqrt{3}}$
458	$\frac{\sqrt{3}}{2}$	344	$\frac{1}{2}$		
678	$\frac{\sqrt{3}}{2}$	355	$\frac{1}{2}$		

Pauli matrices, also obey anti-commutation relations.

$$\lambda_i \lambda_j + \lambda_j \lambda_i = \frac{2}{3}\delta_{ij} + 2 d_{ijk} \lambda_k. \tag{1.21}$$

The non-vanishing values of f_{ijk} (antisymmetric under permutation of any two indices) and d_{ijk} (symmetric) are listed in Table 1.3.

Group operators allow one to move from state to state on the diagrams. The familiar raising and lowering operators $I_\pm = I_x \pm i I_y$ change I_z by one unit, moving left or right on the diagram, and in the larger group there are raising and lowering operators along the 60° lines, called U-spin and V-spin, that change both I_z and Y. The set $(I, I_z), (U, U_z)$, and (V, V_z) has nine operators, so they cannot all be independent. U_+ connects the Σ^+ and the proton, while V_+ connects the Σ^- and the neutron. On the octet, U-spin connects states of the same charge. The axes are shown in Figure 1.7. The definitions in terms of the λ's are:

$$I_\pm = (\lambda_1 \pm i\lambda_2)/2;\ U_\pm = (\lambda_6 \pm i\lambda_7)/2;\ V_\pm = (\lambda_4 \pm i\lambda_5)/2;$$
$$I_z = \lambda_3/2;\ Y = \lambda_8/\sqrt{3}. \tag{1.22}$$

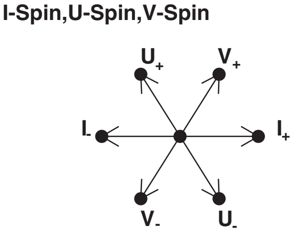


Figure 1.7: I-spin, U-spin, V (all) spin. The operators are not independent. $V_z - U_z = I_z$.

The commutation rules for I-spin, U-spin, and V-spin are the same, namely

$$[I_+, I_-] = 2I_z;\ [U_+, U_-] = 2U_z;\ [V_+, V_-] = 2V_z;\ I_z = V_z - U_z;$$
$$2V_z = \frac{3}{2}Y + I_z;\ 2U_z = \frac{3}{2}Y - I_z. \tag{1.23}$$

Commutation relations between different operators can be worked out using the f_{ijk}'s in Table 1.3. For example,

$$\begin{aligned}[I_z, U_\pm] &= \frac{1}{4}[\lambda_3, \lambda_6 \pm i\lambda_7] \\ &= \frac{1}{4}[\lambda_3, \lambda_6] \pm \frac{i}{4}[\lambda_3, \lambda_7] = -i\frac{\lambda_7}{4} \mp \frac{\lambda_6}{4} = \mp\frac{U_\pm}{2}.\end{aligned} \tag{1.24}$$

Referring to the quarks in Figure 1.2:

$$I_- \left|u\right\rangle = a\left|d\right\rangle;\ \left\langle u\right| I_+ I_- \left|u\right\rangle = \left\langle u\right| [I_+, I_-] \left|u\right\rangle = 2I_z = 1 = |a|^2, \tag{1.25}$$

and

$$V_- \left|u\right\rangle = b\left|s\right\rangle;$$
$$\left\langle u\right| V_+ V_- \left|u\right\rangle = \left\langle u\right| [V_+, V_-] \left|u\right\rangle = 2V_z = \frac{3}{2}Y + I_z = 1 = |b|^2. \tag{1.26}$$

So all three of the matrix elements going around the quark triangle can be chosen +1, in agreement with the convention adopted for I-spin. For the anti-quarks the product of all three is -1, and we adopt a minus sign applied to $\langle \bar{s} | U_+ |\bar{d}\rangle$, in order to preserve the signs of the I-spin matrix elements. With this machinery the matrix elements for the octet can be worked out. It is simpler to work with the meson octet; the matrix elements for the baryon octet are the same. So we construct $3 \otimes 3^* = 8 \oplus 1$ by placing the center of the anti-quark triangle on the corners of the quark triangle, giving a hexagon with three states in the center. Placing the anti-quarks on the u quark gives $u\bar{d}$ at $I_z = +1$, $Y = 0$; $u\bar{s}$ at $I_z = 0$, $Y = 1$, and $u\bar{u}$ at $I_z = 0$, $Y = 0$. Now $U_+ |u\bar{d}\rangle = |u\rangle U_+ |\bar{d}\rangle = -|u\bar{s}\rangle$, so that link has a minus sign, but all of the other operators around the outside of the hexagon have matrix element +1. Complications arise when you jump from the boundary into the middle, where there are three states, that from our construction are $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$. Hop in from the right-hand corner:

$$I_- |u\bar{d}\rangle = |d\bar{d}\rangle + |u\bar{u}\rangle = \sqrt{2}\frac{|d\bar{d}\rangle + |u\bar{u}\rangle}{\sqrt{2}}, \tag{1.27}$$

where we have normalized the wave function, and the matrix element is $\langle \pi^0 | \mathrm{I}_- |\pi^+\rangle = \sqrt{2}$, using the pseudoscalar meson nonet to label the states. I-spin does not couple to the $I = 0$ states, so we have to get there from above or below. We try:

$$\begin{aligned} V_- |K^+\rangle &= V_- |u\bar{s}\rangle = |s\bar{s}\rangle + |u\bar{u}\rangle ; \\ U_+ |\bar{K}^0\rangle &= U_+ |s\bar{d}\rangle = -|s\bar{s}\rangle + |d\bar{d}\rangle . \end{aligned} \tag{1.28}$$

Neither one of these states is orthogonal to $|\pi^0\rangle$. They are linear combinations of the octet member with $I = 0$, called $|\eta\rangle$, and the SU(3) singlet $|\eta'\rangle$. The correct orthogonal forms are

$$\begin{gathered} |\eta\rangle = \frac{1}{\sqrt{6}}(2|s\bar{s}\rangle + |u\bar{u}\rangle - |d\bar{d}\rangle); \quad |\eta'\rangle = \frac{1}{\sqrt{3}}(-|u\bar{u}\rangle + |d\bar{d}\rangle + |s\bar{s}\rangle); \\ I_\pm |\eta\rangle = 0. \end{gathered} \tag{1.29}$$

The octet matrix elements are

$$\langle\eta|\,V_-\,\left|K^+\right\rangle = \frac{1}{\sqrt{6}}(2\,\langle s\bar{s}| + \langle u\bar{u}| - \left\langle d\bar{d}\right|)(|s\bar{s}\rangle + |u\bar{u}\rangle) = \frac{3}{\sqrt{6}} = \sqrt{\frac{3}{2}},$$

and

$$\left\langle\pi^0\right|\,V_-\,\left|K^+\right\rangle = \frac{1}{\sqrt{2}}(\langle u\bar{u}| + \left\langle d\bar{d}\right|)(|s\bar{s}\rangle + |u\bar{u}\rangle) = \frac{1}{\sqrt{2}}. \qquad (1.30)$$

There are 16 elements for the octet, arranged in the following Table 1.4 using the pseudoscalar meson labels. It is important to remember that the matrix elements for the baryon octet are the same, for instance $\left\langle\eta^0\right|\mathrm{V}_-\left|\mathrm{K}^+\right\rangle = \langle\Lambda|\,\mathrm{V}_-\,|\mathrm{p}\rangle$. We shall need these numbers in Chapter 5, in the discussion of hyperon beta decay.

If I-spin symmetry were exact, all particles in a given multiplet would have the same mass. This is not true, but it is close, good to a few percent for the mesons, and better accuracy for the baryons. A good example of the results of this line of reasoning is the Λ hyperon at 1115 MeV. The Σ^+ and Σ^- were known at 1189 and 1197 MeV, respectively, but the Λ was 75 MeV lighter, and hence too far away to be placed in the triplet. A new baryon, the Σ^0 that decayed electromagnetically: $\Sigma^0 \to \Lambda + \gamma$ was predicted, and observed to complete the triplet, and the Λ was assigned $I = 0$. For the Σ triplet one could imagine a symmetry breaking mass term: $m(I_z) = (1193 - 4 \times I_z)$ MeV works very well — try it.

By a similar argument, if unitary symmetry were exact, all of the baryons in the octet would have the same mass. I-spin works better than unitary symmetry. The average nucleon mass is 939 MeV, the Σ is 1193 MeV, and the Ξ is 1318, for an average octet mass of 1150 MeV, and variations of the order of 150 MeV — 15% or so. The meson octet is more difficult to accommodate, with the K meson more than three times as massive as the pion. A suitable symmetry breaking operator for the decuplet is linear in the hypercharge Y:

$$\text{Decuplet } M_{10}(Y) = M_0 + M_1 \times Y, \qquad (1.31)$$

Table 1.4: Matrix elements for the octet of psuedoscalar mesons.

Initial state	Final state	Operator	Matrix element
Around the hexagon			
K^+	π^+	U_-	-1
π^+	$\bar{K}^0$	V_-	1
$\bar{K}^0$	K^-	I_-	1
K^-	π^-	U_+	1
π^-	K^0	V_+	1
K^0	K^+	I_+	1
Jumps to the middle			
π^+	π^0	I_-	$\sqrt{2}$
π^0	π^-	I_-	$\sqrt{2}$
K^+	π^0	V_-	$\frac{1}{\sqrt{2}}$
K^0	π^0	U_-	$-\frac{1}{\sqrt{2}}$
$\bar{K}^0$	π^0	U_+	$\frac{1}{\sqrt{2}}$
K^-	π^0	V_+	$\frac{1}{\sqrt{2}}$
K^+	η^0	V_-	$\sqrt{\frac{3}{2}}$
K^0	η^0	U_-	$\sqrt{\frac{3}{2}}$
$\bar{K}^0$	η^0	U_+	$-\sqrt{\frac{3}{2}}$
K^-	η^0	V_+	$\sqrt{\frac{3}{2}}$

where $M_0 = m_{\Sigma^*} = 1385\,\text{MeV}$, and $M_1 = -149\,\text{MeV}$, predicts $m_\Delta = 1236\,\text{MeV}$, and $m_{\Xi^*} = 1534\,\text{MeV}$, good to better than 1%. This 'equal spacing rule' predicts a new particle at the bottom of the triangle with $S = -3$, $I = 0$, and mass of about 1680 MeV. The hypothetical baryon is named the Ω^-, and its strangeness conserving decay $\Omega^- \to \Xi^0 + K^-$ is energy forbidden, so unlike all of its partners in the decuplet, the Ω^- should decay by weak interaction.

Table 1.5: Baryon octet masses from Equation (1.33).

Baryon	I	Y	$(I(I+1)-Y^2/4)$	M_8(MeV)	Book value
p, n	1/2	1	1/2	943	939
Σ	1	0	2	1187	1193
Λ	0	0	0	1115	1115
Ξ	1/2	−1	1/2	1323	1318

Soon after its prediction the Ω^- was observed in 1964 at Brookhaven with a mass of 1672 MeV, a lifetime of 0.8×10^{-10} s, and all of the expected properties [6].

Equation (1.31) won't work for the baryon octet, because it does not allow any mass splitting between the Λ and the Σ^0 (77 MeV), so a more elaborate operator was constructed that involves I spin as well as hypercharge:

$$\text{Octet } M_8(I,Y) = M_0 + M_1 \times Y + M_2 \times (I(I+1) - Y^2/4). \quad (1.32)$$

The parameters M_0, etc., are representation dependent, and not necessarily the same in Equations (1.31) and (1.32). For the decuplet, set $M_2 = 0$. An illustrative, if not definitive, set of numbers for the octet is given in Equation (1.33):

$$M_8(I,Y) = 1115 - 190 \times Y + 36 \times (I(I+1) - Y^2/4)\,\text{MeV}. \quad (1.33)$$

This formula, chosen to match the Λ mass exactly, gives the numbers in Table 1.5.

The formula is good to ~1/2 %. An alternate combination from Equation (1.32) is the Gell-Mann–Okubo mass formula:

$$3M_\Lambda/4 + M_\Sigma/4 = (M_n + M_\Xi)/2. \quad (1.34)$$

This good agreement on the mass splitting, plus the prediction of the Ω^- at the bottom of the decuplet, gave support to the SU(3) classification of the hadrons. Evidence for the quarks as elementary substructure of the hadrons was hindered by the absence of free quarks in the laboratory, but if the quark model is accepted, the hadrons lose their elementary particle status, making Table 1.6 obsolete.

Table 1.6: Hadrons of Table 1.1 in the quark model.

	Mass (MeV)	Spin/ parity	I-spin	I_z	Baryon number	Strangeness
Quarks symbol						
u	300	$1/2^+$	1/2	$+1/2$	1/3	0
d	300	$1/2^+$	1/2	$-1/2$	1/3	0
s	500	$1/2^+$	0	0	1/3	-1
Antiquarks						
$\bar{s}$	500	$1/2^+$	0	0	$-1/3$	$+1$
$\bar{d}$	300	$1/2^+$	1/2	$+1/2$	$-1/3$	0
$\bar{u}$	300	$1/2^+$	1/2	$-1/2$	$-1/3$	0

[*] These are approximate 'constituent quark' masses $m_u = m_d \sim m_p/3$, while $2m_u + m_s \sim m_\Lambda$. A spin dependent term has to be added to a simple sum to fit both meson and baryon masses. The necessity for such a term is clear from the pseudoscalar–vector meson octet comparison. They have the same quark content, but the vectors are heavier than the pseudoscalars.

Name	Symbol	Mass (MeV)	Spin/ parity	I-spin	Strangeness	Quark content
Pseudoscalar meson nonet, spin zero						
Pion	$\pi^\pm$	139.6	0^-	1	0	$u\bar{d};\ d\bar{u}$
	π^0	135.0	0^-	1	0	$(d\bar{d} + u\bar{u})/\sqrt{2}$
Eta	η	548	0^-	0	0	$(-d\bar{d} + u\bar{u} + 2s\bar{s})/\sqrt{6}$
Eta′	η'	958	0^-	0	0	$(d\bar{d} - u\bar{u} + s\bar{s})/\sqrt{3}$
Kaon	K^+	494	0^-	1/2	$+1$	$u\bar{s}$
	K^0	498	0^-	1/2	$+1$	$d\bar{s}$
	$\bar{K}^0$	498	0^-	1/2	-1	$s\bar{d}$
	K^-	494	0^-	1/2	-1	$s\bar{u}$
Vector meson nonet, spin 1						
Rho	$\rho^\pm$	773.5	$1-$	1	0	$u\bar{d}; d\bar{u}$
	ρ^0	773.5	$1-$	1	0	$(d\bar{d} + u\bar{u})/\sqrt{2}$
Omega	ω	782.6	$1-$	0	0	$(-d\bar{d} + u\bar{u} + 2s\bar{s})/\sqrt{6}$
Phi	ϕ	1020	$1-$	0	0	$(d\bar{d} - u\bar{u} + s\bar{s})/\sqrt{3}$
K^*	K^*	892	$1-$	1/2	$+1$	$u\bar{s}; d\bar{s}$
$\bar{\mathrm{K}}^*$		892	$1-$	1/2	-1	$s\bar{u}; s\bar{d}$

(*Continued*)

(*Continued*)

Name	Symbol	Mass (MeV)	Spin/ parity	I-spin	Strangeness	Quark content
Baryons spin 1/2						
Proton	p	938.3	$1/2^+$	1/2	0	(uud)
Neutron	n	939.6	$1/2^+$	1/2	0	(udd)
Lambda	Λ	1115.6	$1/2^+$	0	−1	(uds)
Sigma	Σ^+	1189	$1/2^+$	1	−1	(uus)
	Σ^0	1192.6	$1/2^+$	1	−1	(uds)
	Σ^-	1197.4	$1/2^+$	1	−1	(dds)
Cascade	Ξ^0	1315	$1/2^+$	1/2	−2	(uss)
	Ξ^-	1322	$1/2^+$	1/2	−2	(dss)
Baryons spin 3/2						
Delta	Δ^{++}	1232	$3/2^+$	3/2	0	(uuu)
Delta	Δ^+	1232	$3/2^+$	3/2	0	(uud)
Delta	Δ^0	1232	$3/2^+$	3/2	0	(udd)
Delta	Δ^-	1232	$3/2^+$	3/2	0	(ddd)
Sigma∗	Σ^{*+}	1385	$3/2^+$	1	−1	(uus)
Sigma∗	Σ^{*0}	1385	$3/2^+$	1	−1	(uds)
Sigma∗	Σ^{*-}	1385	$3/2^+$	1	−1	(dds)
Cascade∗	Ξ^{*0}	1530	$3/2^+$	1/2	−2	(uss)
Cascade∗	Ξ^{*-}	1530	$3/2^+$	1/2	−2	(dss)
Omega	Ω^-	1672	$3/2^+$	0	−3	(sss)

The meson nonet quark wave functions display SU(3) Clebsch–Gordan coefficients for $3 \otimes 3^* = 8 \oplus 1$, and the matrix elements connecting the octet states are given in Table 1.4. The SU(3) singlet η' can be reached by the U and V operators, and the matrix elements can be calculated in the same way. The two mesons in the middle, η and η' for the pseudoscalars and ω and ϕ for the vectors can mix, because they have the same quantum numbers, and the SU(3) assignments are arbitrary. This is especially important for the ϕ, which decays predominantly $\phi \to K + \bar{K}$, and must have quark content $(s, \bar{s})$. A mixing angle $\theta = -35°$, where $|\phi'\rangle = \cos(\theta)|\phi\rangle + \sin(\theta)|\omega\rangle$ does the trick, and we shall henceforth assume that $\phi = (s, \bar{s})$.

For the baryon octet, the flavors (u, d, s) are correct, but the SU(3) wave functions take into account permutations of the quarks from the three different 3's, and the fact that there are two different octets with orthogonal wave functions. The Λ and the Σ^0 are both (uds), but the quark wave functions are orthogonal. Color symmetry can be helpful here (see below). All hadrons must be color singlets, and two like flavored quarks, like (uu) in the proton and the Σ^+ must be a color 3*, so that the 3 of the extra quark can form a singlet. The color 3* is totally antisymmetric, so the two u quarks must be in a triplet spin state, like a deuteron. It follows that the Σ^0 must have a triplet spin (ud) pair, so that the Λ has a singlet spin (ud) pair, and is orthogonal. The spin of the Λ is the spin of the strange quark.

1.7. Color

In the quark model the decuplet has three states with three identical quarks, all with parallel spins: Δ^{++} (u, u, u), Δ^- (d, d, d), and Ω^- (s, s, s). Each quark has spin 1/2, and should obey the Pauli exclusion principal, so two identical quarks cannot occupy the same state. This apparent contradiction was solved by the hypothesis of another SU(3) group called color. Each quark comes in three colors, red, blue and yellow, say, and the Δ^{++} has one u quark of each color, so they are not identical! Color is assumed to be a quantum number that is confined, that is to say that observable hadrons are color singlets. In the decomposition $3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1$, the 10 is totally symmetric, the 8's are mixed, and the singlet is totally antisymmetric. The colored quarks in the baryon decuplet are in the singlet color state, because the baryons have no color. Similarly for the mesons the color decomposition of $3 \otimes 3^* = 8 \oplus 1$, and the observable mesons are color singlets. Color is an important symmetry for QCD, and will appear later in the book. Massless spin-one gluons are the analog of the photon in QCD, and are responsible for the binding force between the quarks. Gluons carry no electric charge, but do carry color charges. For example a red u quark and a blue u quark would interact by the exchange of a $\mathrm{R\bar{B}}$ or $\mathrm{B\bar{R}}$ gluon.

1.8. Applications of Unitary Symmetry

The strange mesons and baryons dominantly decay by weak interaction into hadrons. However, they also decay leptonically: $K^+ \to \mu^+ + \nu_\mu$ (the analog of pion decay), and semileptonically: $K^+ \to \pi^0 + \mu^+ + \nu_\mu$. The Λ, Σ, and Ξ hyperons also decay semileptonically, like the neutron: $n \to p + e^- + \bar{\nu_e}$. With sufficient energy available in the hyperon decays muons can replace electrons, and both decays occur. The two body decay $K^+ \to e^+ + \nu_e$ is strongly suppressed by spin selection rules. Applying SU(3) symmetry to baryon decay predicts matrix elements $\langle p | I_+ | n \rangle = 1$; $\langle p | V_+ | \Lambda \rangle = \sqrt{3/2}$; $\langle n | V_+ | \Sigma^- \rangle = 1$, and $\langle \Sigma^+ | \mathrm{Octet} | \mathrm{n} \rangle = 0$, where Octet is any octet operator. The beta decay $\Sigma^+ \to n + \mu^+ + \nu_\mu$ is forbidden, while $\Sigma^- \to n + \mu^- + \bar{\nu_\mu}$ is allowed, consistent with experiment.[g] This is one example of the $\Delta Q = \Delta S$ rule, satisfied by the semileptonic weak decays, and explained by unitary symmetry. Another example from the mesons is that $\left\langle K^0 \right| H_W \left| \pi^- \right\rangle$ and $\left\langle \bar{K^0} \right| H_W \left| \pi^+ \right\rangle$ are OK, but not with the pions interchanged.

The $|\Delta I| = 1/2$ rule also applies to strange particle decays. For the K mesons a final state is $K \to \pi + \pi$. Each pion has $I = 1$, so a pair of pions can have $I = 0, 1, 2$. However, the $\pi^0\pi^0$ pair can have only $I = 0$ or 2, because the relevant Clebsch–Gordan coefficient vanishes (see Table 1.2). So assuming the I-spin assignment is the same for all final states, only $I = 0$ or $I = 2$ are allowed. Since the K meson has $I = 1/2$, $|\Delta I| = 1/2$ requires the $K \to \pi + \pi$ final state to have $I = 0$. But $K^+ \to \pi^+ + \pi^0$ cannot have $I = 0$, because it is positively charged, so $K^+ \to \pi^+ + \pi^0$ violates the rule. The relevant lifetime comparison in Table 1.1 is $\tau_{K^+}/\tau_{K_s^0} = 135$, showing that in this instance $|A(\Delta I = 3/2)|/|A(\Delta I = 1/2)| = 0.09$.

1.9. The Neutral *K* Mesons

The neutral K meson system has many interesting properties. Strangeness or hypercharge assignments give the K^0 and $\bar{K^0}$ different quantum numbers that act like charge for strong interactions.

---

[g] $(\Sigma^+ \to n + e^+ + \bar{\nu_e})/(\Sigma^- \to n + e^- + \bar{\nu_e}) < 0.043$ (see Ref. [7]).

We shall learn in Chapter 3 that a charged field has two components with the same mass. Hypercharge plays this role for the K's. The strong interaction produces K^0 and $\bar{K^0}$, a particle–antiparticle pair. But the weak interaction allows both to decay into two pions: $K^0 \to \pi^+ + \pi^-$, and $\bar{K^0} \to \pi^+ + \pi^-$, so the weak interaction mixes particle and antiparticle. This same phenomenon occurs for neutral mesons made of heavier quarks, called $D^0(\bar{u}, c)$ and $B^0(\bar{b}, d)$, where we have introduced new quarks called charm c with $e_q = +2/3$, and bottom b with $e_q = -1/3$. These quarks carry new conserved quantum numbers like strangeness S: $C = +1$ for the c quark, and $B = -1$ for the b quark. Note that the B mesons have the $\bar{b}$ quark, like the K mesons with the $\bar{s}$ quark, while the D meson contains the c quark — not the $\bar{c}$. Discussion of the heavy quark mixing will be deferred to a later chapter.

Two new particles with definite lifetimes, called K_S and K_L for short and long lifetimes, respectively (see Table 1.1) were introduced to describe the mixing of K^0 and $\bar{K^0}$.

$$|K^0\rangle = (|K_S\rangle + |K_L\rangle)/\sqrt{2}, \tag{1.35}$$

$$|\bar{K^0}\rangle = (|K_S\rangle - |K_L\rangle)/\sqrt{2}. \tag{1.36}$$

It is the K_S that decays into two pions. At this point we introduce some discreet symmetries that lead to conserved quantities. Charge Q is one example, and baryon number N is another. No reaction exists that does not conserve charge, and nature seems also to respect baryon number. The charge conjugation operator, C, interchanges particle and antiparticle. $C^2 = 1$. For the pions $C|\pi^+\rangle = -|\pi^-\rangle$, and $C|\pi^0\rangle = +|\pi^0\rangle$ (see the G parity Problem 1.4). For the K mesons $C|K^+\rangle = |K^-\rangle$, and $C|K^0\rangle = |\bar{K^0}\rangle$. The antiparticle to the Σ^+ is the $\bar{\Sigma}^-$, part of the antibaryon octet. C reverses all quantum numbers of the particle, I_Z, Q, S, and N. If the particle has intrinsic spin, S_Z is also reversed. Charge conjugation invariance is respected by the strong and electromagnetic interactions. The photon is odd under charge conjugation $C|\gamma\rangle = -|\gamma\rangle$, because the coupling to the electromagnetic field is proportional to the charge. This guarantees $C|\pi^0\rangle = +|\pi^0\rangle$.

Parity is a quantum number familiar from atomic physics. Wave functions are even or odd under reflection of the space coordinates through the origin. Hydrogen wave functions have the parity of l, the orbital quantum number. Hydrogen energy levels depend on the principal quantum number n, with $0 \leq l \leq n-1$. The Hamiltonian contains the L^2 operator, so the energy levels could be l dependent, but they are not. This l degeneracy is a property of the Coulomb potential. Parity is adopted in elementary particle physics. Particles are assigned a relative intrinsic parity; pions and kaons are pseudoscalars with spin zero and odd intrinsic parity. The octet baryons have spin 1/2 and are assigned even parity. Odd relative parity in a meson–baryon system is assigned to the meson. Parity is conserved by the strong and electromagnetic interactions. A clean demonstration of the odd intrinsic parity of the pion was the measurement of the capture of π^- in deuterium: $\pi^- + d \rightarrow n + n$. The Pauli principle requires the two identical fermions in the final state — the two neutrons — to have $L_J^{2s+1} = S_0^1, P_{0,1,2}^3, D_2^1, F_{2,3,4}^3$, etc., spin zero and even orbital angular momentum; spin one and odd orbital angular momentum. The states have the parity of l. Capture of the π^- occurs mainly from $l = 0$ s states where the amplitude of the pion wave function is large at the nucleus. The deuteron has spin-parity 1^+, so if the pion has odd intrinsic parity, the $\pi^- + d$ initial state has J^P: S wave 1^-; P wave $2^+, 1^+, 0^+$; D wave $3^-, 2^-, 1^-$. S wave matches the P_1^3 two neutron odd parity final state. So the observation of the capture reaction with reasonable probability establishes the odd intrinsic parity for the π^- [8]. It is natural to assume that the π^+ has odd parity as well.

The parity of the π^0 can be defined independent of interaction with baryons, because of the electromagnetic decay $\pi^0 \rightarrow \gamma + \gamma$. First of all, this decay precludes spin one. That a particle with spin one cannot decay into two photons was first proven by Landau [9]. Electromagnetic interactions conserve parity, and the parity of the final state is correlated with the polarizations of the outgoing photons. Even parity gives two photons with parallel electric vectors, while odd parity gives perpendicular electric vectors. Linear polarization states for 67 MeV photons are difficult to measure (although circular polarization states can be measured). However, when the photon converts spontaneously into an e^-e^+ pair, (called a Dalitz pair after R. H. Dalitz) the electric vector tends to lie in the plane of the pair. Odd parity for the π^0 was established from the angular correlation of the planes of e^+e^- pairs in the rare decay $\pi^0 \to 2e^+ + 2e^-$ [10].

The relative parity of the K^- was established by studying K^- capture in helium: $K^- + \mathrm{He}^4 \to \pi^- + \mathrm{He}^4_\Lambda$. He^4_Λ is called a hyperfragment, where one of the two neutrons in the He nucleus is replaced by a Λ hyperon, so $\mathrm{He}^4_\Lambda = p, p, n, \Lambda$. Just like the He^4 nucleus, the hyperfragment has spin zero. Since the kaon and the pion also have spin zero, there is no spin in the problem, and the the parity of the K^- must be the same as the π^- [11]. The total binding energy of the He^4 nucleus is only 28 MeV, while the pion momentum in the capture reaction: $K^- + n \to \pi^- + \Lambda$ is 254 MeV/c, so it is remarkable that some of the time the Λ is left sitting in the helium nucleus where the neutron was.

The weak interactions do not conserve parity, so the decay final states are not useful to fix the intrinsic parity of the parent. Theorists assumed, however that the product CP of charge conjugation times parity was still conserved. In the K rest frame the $(\pi^+ + \pi^-)$ final state is CP even. C interchanges the pions, and P restores them to their original position. The orbital angular momentum $l = 0$. The state $|K^0_S\rangle$ in Equation (1.23) is CP even: $\mathrm{CP}|K^0_S\rangle = +|K^0_S\rangle$, and $|K^0_L\rangle$ is CP odd: $\mathrm{CP}|K^0_L\rangle = -|K^0_L\rangle$. $K^0_S \to \pi + \pi$, while this final state is CP forbidden for K^0_L. This assignment is consistent with $\mathrm{CP}|\mathrm{K}^0\rangle = |\bar{\mathrm{K}^0}\rangle$. $K^0_L \to \pi\pi\pi$ is allowed. Both K^0_S and K^0_L can decay semileptonically into $\pi\mu\nu$ or $\pi e\nu$. This analysis has to be modified (complicated) by the addition of the small probability that K^0_L does indeed decay into two pions — the celebrated CP violation, that will be discussed in Chapter 9.

Table 1.1 shows that the K^0_S and K^0_L have different lifetimes τ_S and τ_L. The time dependent K^0 state vector can be written apart from an overall phase as

$$|K^0(t)\rangle = (e^{-t/(2\tau_s)} \times |K^0_S\rangle + e^{-t/(2\tau_L)} \times |K^0_L\rangle e^{-i\Delta m t})/\sqrt{2}, \quad (1.37)$$

where we have allowed for a mass difference between K^0_S and K^0_L. Time is measured in the K^0 rest frame. Substituting K^0 and $\bar{K^0}$ for K^0_S and K^0_L, and calculating the square of the amplitude $|\langle \bar{K_0}|K_0\rangle|^2$

gives the equation:

$$|\langle \bar{K_0}|K_0\rangle|^2 = (e^{-t/\tau_s} + e^{-t/\tau_L} - 2\cos(\Delta m t)e^{-(1/2\tau_s + 1/2\tau_L)t})/4. \tag{1.38}$$

Note that this expression vanishes at $t = 0$, but does not vanish for finite t. So as the K^0 moves away from the production vertex a finite amplitude for the strangeness -1 $\bar{K}^0$ appears, and this amplitude can produce Λ and Σ hyperons. This phenomenon is called strangeness oscillation, and depends on the magnitude, but not the sign, of the mass difference.

The two component wave functions in Equations (1.35) and (1.36) give rise to many interesting phenomena. The $K_L - K_S$ mass difference has been measured. It is very tiny, $\sim 10^{-5}$ eV, and its sign is known to be positive, i.e. the K_L is heavier. For oscillation $\Delta m \times t \sim 1$ implies $t \sim 10^{-10}$ s, the K_S lifetime.

Half of the K^0 produced in $\pi^- + p \to K^0 + \Lambda$ are K_S and decay near the production vertex, and half are K_L and travel several meters. If you place a target several meters downstream, because the $\bar{K_0}$ component of the K_L interacts more with matter than the K^0, the balance is altered, and K_S reappear in the beam! This is called regeneration. The CP violation, allowing $K_L \to \pi + \pi$, enriches this phenomenology. It is a good example of wave–particle duality and is discussed in detail in Chapter 9.

1.10. Electron–Positron Storage Rings

A storage ring is a synchrotron in which the beam, after reaching full energy, coasts indefinitely. Beam lifetimes can be hours, even days, depending on the quality of the ring. In 1960, Bruno Touschek at the Italian National Laboratory in Frascati proposed injecting positrons and electrons counter rotating in the same ring. It took a few years to realize Touschek's idea — intense positron sources suitable for injection had to be developed, for one thing. But once proven practical, these rings have been enormously productive in the study of elementary particle physics. Although some asymmetric pairs of rings have been built for special effects, most particle — antiparticle colliders are single rings with counter rotating beams of the same energy. If E is the energy of the ring, the total energy available in the center of mass is $2E$.

1.11. Luminosity

Another important parameter is the luminosity:

$$\mathcal{L} = \frac{f \times n_1 \times n_2}{A}. \tag{1.39}$$

The units of luminosity are $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, and the event rate for a cross-section σ is $w = \sigma \times \mathcal{L}$. In Equation (1.39), f is the bunch crossing frequency, n_1 and n_2 are the number of particles in each bunch, and A is the transverse area overlap. If there are N bunches of each beam in the machine, multiply $\mathcal{L}$ by N. In an e^-e^+ or $p\bar{p}$ collider there are usually more particles than antiparticles. The frequency depends on the circumference of the ring. The SPEAR ring at SLAC is 80 m in diameter, giving $f = 1.2\,\mathrm{MHz}$ (beam velocity is c). Beams in e^-e^+ colliders are typically ribbon shaped in cross-section, with height $10\,\mu m$ and width $100\,\mu m$ for an area A $\sim 10^{-5}\mathrm{cm}^2$. If $n_1 = 10^{11}$ and $n_2 = 3 \times 10^{10}$ particles, then $\mathcal{L} = 3.6 \times 10^{32}\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, a reasonable number. Integrated luminosity $\int \mathcal{L}dt$ is usually quoted in inverse picobarns $-1\ \mathrm{pb}^{-1} = 10^{36}\,\mathrm{cm}^{-2}$, which would be about 45 min at $\mathcal{L} = 3.6 \times 10^{32}$. If a process has a cross section of 10 pb, on average one would have 10 events in the data sample for 100% detection efficiency.

1.12. Natural Units

Maxwell's equations are usually written in mks units, with charge in coulombs, current in amperes = coulombs/s, electric fields in volts/m, and magnetic fields in tesla. Volts times amperes = watts. In these units the equations are

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0}, \tag{1.40}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \tag{1.41}$$

$$\vec{\nabla} \cdot \vec{B} = 0, \tag{1.42}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \cdot \vec{j} + \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t}, \tag{1.43}$$

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}). \tag{1.44}$$

The constants $\epsilon_0 = 8.85 \times 10^{-12}, \mu_0 = 4\pi \times 10^{-7}$, giving $1/(\mu_0 \epsilon_0) = c^2 = 9 \times 10^{16} (\mathrm{m/s})^2$. Coulomb's law in these units reads:

$$V(r) = \frac{q_1 q_2}{4\pi\epsilon_0 r}. \tag{1.45}$$

This form is not popular in the Schrödinger equation, where V(r) = $q_1 q_2/r$ is preferred. That corresponds to Maxwell's equations in cgs units, with charge measured in electrostatic units. A factor of 4π appears on the right hand side of the first Maxwell Equation, and there is no ϵ_0. Electrostatic units have certain advantages. The E and B fields have the same dimensions, as do the electric and magnetic dipole moments. The elementary charge $e = 1.602 \times 10^{-19}$ coulombs = 4.8×10^{-10} esu. Planck's constant in atomic physics is in cgs units, or more conveniently in eV-s: $\hbar = 6.58 \times 10^{-16}$ eV-s, and the speed of light is $c = 3 \times 10^{10}$ cm/s. A useful product is $hc = 1240$ eV-nm, or if you prefer $\hbar c = 197.3$ eV-nm. The fine structure constant in cgs units is $\alpha = e^2/\hbar c$, and in mks units $\alpha = e^2/(4\pi\epsilon_0 \hbar c)$, but the numerical value of this dimensionless constant is always the same: $\alpha = 1/137.035$.

In particle physics we are accustomed to measuring energy in MeV, momentum in MeV/c, and mass in MeV/c^2. Natural units set $\hbar = c = 1$. In these units energy, momentum, and mass have the same dimensions — MeV, or GeV = 1000 MeV. Length and time have the same dimensions, and energy 1/length. The Compton wave length of a particle $\lambda_\mathrm{c}/2\pi = \hbar/\mathrm{mc}$ becomes simply $\lambda_\mathrm{c}/2\pi = 1/\mathrm{m}$. The expression $\hbar/mc$ is a length, and $\hbar/mc^2$ is a time, and in natural units both are 1/m. So if you are calculating a length or a time, you have to convert the result to get cm or s. Evaluate $(\hbar/mc^2) \times c$, with $mc^2 = 1000$ MeV and square the result to obtain $1(\mathrm{GeV})^{-2} = 0.389$ mb $= 0.389 \times 10^{-27}\,\mathrm{cm}^2$. This is a very useful conversion constant. The fine structure constant is $\alpha = \mathrm{e}^2/4\pi = 1/137$, so the dimensionless elementary charge in these units is $e = 0.3$.

1.13. Invariant Kinematic Variables

Since E, p, and m all have the same dimensions, a 2 body → 2 body reaction can be conveniently described in terms of three invariant variables. There are four equations to conserve energy and momentum:

$$p_1 + p_2 = p_3 + p_4. \tag{1.46}$$

One component is energy conservation:

$$E_1 + E_2 = E_3 + E_4. \tag{1.47}$$

And three components conserve momentum:

$$\vec{p_1} + \vec{p_2} = \vec{p_3} + \vec{p_4}. \tag{1.48}$$

In our metric the square of a particle momentum four vector is positive:

$$p_1^2 = E_1^2 - \vec{p_1^2} = m_1^2. \tag{1.49}$$

Now introduce the three invariant variables, called s, t, and u:

$$s = (p_1 + p_2)^2 = (E_1 + E_2)^2 - (\vec{p_1} + \vec{p_2})^2, \tag{1.50}$$

$$t = (p_1 - p_3)^2, \tag{1.51}$$

$$u = (p_1 - p_4)^2, \tag{1.52}$$

$$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2, \tag{1.53}$$

where s is the square of the total energy in the center of mass, where $\vec{p_1} + \vec{p_2} = 0$, t is called the 1 → 3 four momentum transfer, and u is the 1 → 4 four momentum transfer. In our metric normally $s > 0, t < 0$, and $u < 0$. These variables are widely used in particle physics. They are sometimes called the Mandelstam variables, after the theorist who introduced them. More about special relativity is covered in Chapter 2.

1.14. Basic e^-e^+ Reactions

Two fundamental QED reactions in an e^-e^+ storage ring are elastic scattering, called Bhabha scattering:

$$e^- + e^+ \to e^- + e^+, \tag{1.54}$$

and muon pair production

$$e^- + e^+ \to \mu^- + \mu^+. \tag{1.55}$$

The one photon exchange diagrams for these two processes are shown in Figure 1.8. The upper diagram is called 't channel', and the lower one, the annihilation diagram, is called 's channel'. The positron arrows are reversed. Time increases from left to right. Bhabha scattering proceeds by both diagrams, giving two quadratic terms and an interference term and a rather complicated expression. Small angle Bhabha scattering has a large cross-section, and is useful as a luminosity monitor in storage rings. Muon pair production has only the s channel annihilation diagram, because there is no mechanism for changing an electron into a muon at a single vertex: $\mu^- \to e^- + \gamma$, although allowed by the conservation of energy,

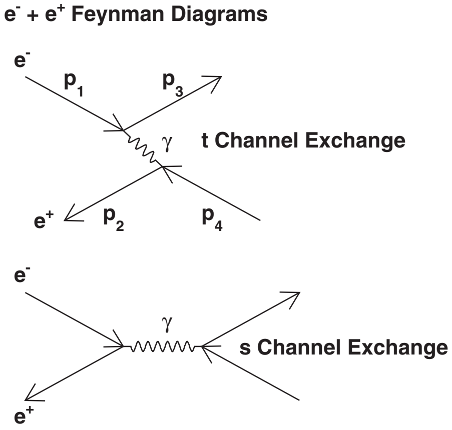


Figure 1.8: First order diagrams in quantum electrodynamics. Both diagrams contribute to Bhabha scattering, but only the s channel diagram contributes to $e^- + e^+ \to \mu^- + \mu^+$.

does not occur. The s channel diagram contains an intermediate virtual photon, that has only energy, and no momentum, but has photon quantum numbers: spin 1, odd under parity, and odd under charge conjugation. QED gives a simple formula for the cross section of unpolarized beams, neglecting the lepton masses:

$$d\sigma/d\Omega = \frac{\alpha^2}{4s}(1+\cos^2\theta), \tag{1.56}$$

where θ is the polar angle of the μ^+ relative to the axis of the colliding beams; $\theta = 0$ is by convention in the positron direction. Integrating over the solid angle $d\Omega = \sin\theta d\theta d\phi$ gives the total cross-section:

$$\sigma(e^- + e^+ \to \mu^- + \mu^+) = \frac{4\pi\alpha^2}{3s}. \tag{1.57}$$

For $s = 10\,\mathrm{GeV}^2$ this formula gives a cross-section of $\sigma = 8.7\times10^{-33}\mathrm{cm}^2$, or 8.7 nb. A luminosity of 3.6×10^{32} gives about 3 events/s. This process is called one unit of R, and is used to normalize e^-e^+ runs. We shall learn how to calculate Equation (1.57) in Chapter 4, where very interesting polarization phenomena will also be discussed.

1.15. e^-e^+ Annihilation and the Quark Model

Equation (1.45) can be applied to quarks:

$$\sigma(e^-e^+ \to q+\bar{q}) = \frac{4\pi\alpha^2}{s}\sum_{e_q} e_q^2 = 3\sum_{e_q} e_q^2 \times R, \tag{1.58}$$

where e_q is the quark charge in units of the elementary charge, and the multiplier color factor is 3. The (u, d, s) quark charges are $+2/3$, $-1/3$, $-1/3$, so the sum $\sum e_q^2 = 2/3$, and the quark model predicts $\sigma(e^- + e^+ \to \text{hadrons}) = 2\times R$. The quark–antiquark pair is not observed in the final state. Rather by a not completely understood process called hadronization or fragmentation, the $(q\bar{q})$ pair acquires other opposite pairs as the quarks fly apart, forming mesons. At low energies $e^-e^+ \to \pi^-\pi^+$, or $\pi^0\pi^0$. At higher energies more hadrons,

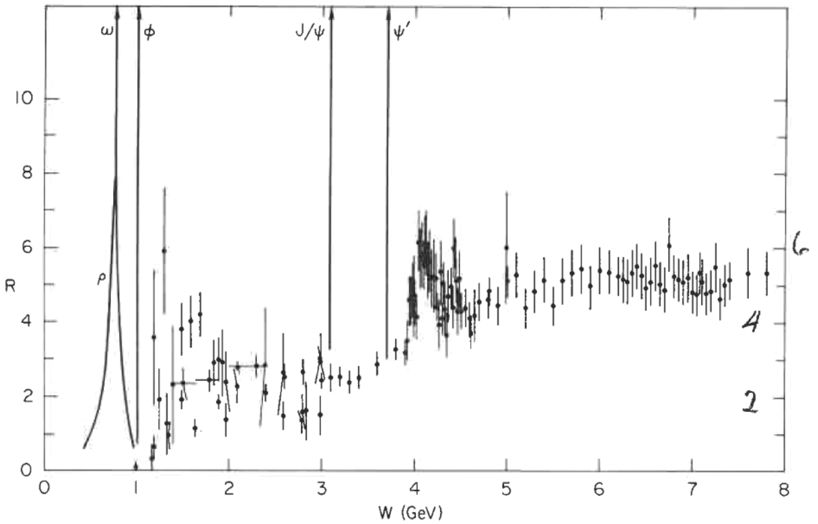


Figure 1.9: $e^-e^+ \to$ hadrons as a function of center of mass energy from data available in 1976.

including strange particles and baryons are produced, giving rise to ‘jets’ of particles streaming out back to back from the production vertex. The jet direction is assumed to reflect the direction of the primary quark. Dividing the cross-section by R removes the steep $1/E^2$ energy dependence.

Figure 1.9 [12, pp. 89–149, Fig. 21] shows the cross-section in units of R for hadron production in e^-e^+ annihilation from $s = 0$ to 8 GeV. The plot is very busy. Near 1 GeV are peaks from the ρ, ω, and ϕ vector mesons. There is still structure between 1 and 3 GeV, but an eye favorable towards the quark model might spot an average of 2. Then at 3.1 GeV there is an explosion, shown on an expanded energy scale in Figure 1.10 [12, Fig. 22]. At 3.05 GeV $R = 9.3$ nb, and the data point is at about 22 nb (it is a log scale), or $R = 2.4$. Before 3.10 GeV the cross-section rises almost vertically, increasing about a factor of 100. There is about 1 data point per 600 keV step through the resonance, with an energy setting error quoted at ±100 keV, or 1 part in 15,000 of the ring energy. The stability and calibration of the ring energy is impressive. These data were taken at SLAC.

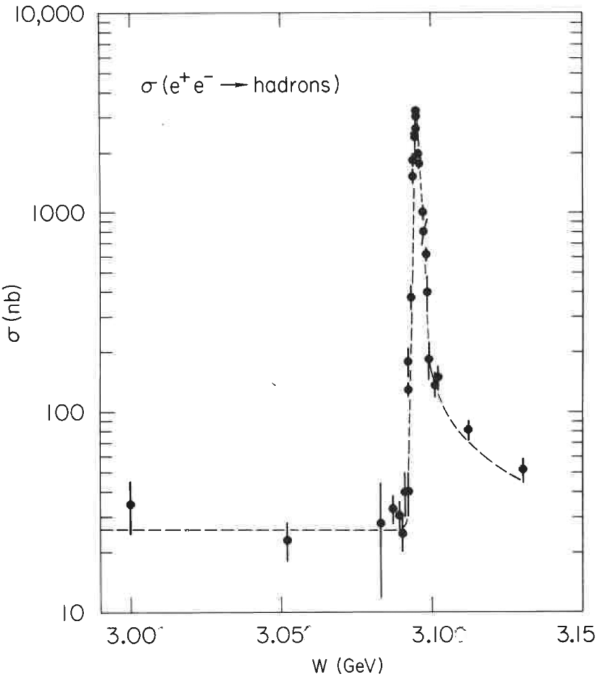


Figure 1.10: Careful energy scan in the region of 3.095 GeV.

The resonance is so narrow — the full width is 93 ± 2 keV — that it is easy to miss if you don't know that it is there.[h] Only later was the entire mass range scanned in 600 keV steps. It is difficult to see the measured line width from the sharp edge in Figure 1.10, but the original paper by Boyarski *et al.* shows an expanded scale, where the peak is centered at 3.095 GeV, with full width of 5 MeV — dominated by radiation and other experimental effects [13]. The shoulder on the high mass side is the result of initial state radiation. Above the resonance the e^-e^+ beams can emit a photon prior to collision, and fall down to the resonance energy. The probability of this happening is a few %, but the resonance is so strong that the cross-section noticeably increases.

---

[h] The line width can be obtained from the integral under the curve even though the observed width is dominated by other effects.

This resonance was independently discovered in a fixed target experiment at Brookhaven National Laboratory where a double arm spectrometer measured the momenta of electron–positron pairs from proton bombardment of a beryllium target. The SLAC group named the new particle 'ψ', while the Brookhaven group named it 'J', so it is now called the J/ψ, with the distinction of being the only elementary particle with two names. The Brookhaven group did not find any higher resonances.[i]

So what is this amazing peak? Boyarski *et al.* establish the J^{PC} of the state as 1^{--}, just like the photon, but make no other comments on its identity. Shortly after the discovery of the $J/\psi(3095)$, a higher mass resonance was found, shown in Figure 1.9, called the $\psi'(3684)$. While the J/ψ has a long list of hadronic and leptonic final states, the ψ' decays predominantly into $J/\psi + \pi^+\pi^-$, $J/\psi + \pi^0\pi^0$, or $J/\psi + \gamma\gamma$, where the third decay proceeds through an intermediate state.

One clue to the identity of the $J/\psi(3095)$ is its very narrow width 90 keV, compared to the $\phi(s, \bar{s})$ meson with a lower mass, and a width of 4.4 MeV. Generally, mesons with higher mass would be expected to have broader widths — i.e. shorter lifetimes. But the J/ψ lives 10^{-20} s, which is like forever for a strong decay. If it were a $(q, \bar{q})$ pair of a new heavy quark, with new quantum numbers like strangeness, and the mass were below threshold to produce mesons carrying this quantum number, then the decay would be strongly inhibited. To first order it would have no place to go.

Contrast this with the ϕ, which is above $K\bar{K}$ threshold, so it can accommodate its strange quarks directly into mesons. One might ask whether such effects below threshold have been observed for $(s, \bar{s})$? The answer is no, and the explanation lies in our understanding of light quark binding. The new quarks in the bound state J/ψ are heavy, with a mass ~ 1.3 GeV, so that the non-relativistic Schrödinger equation can be used to describe the bound states, while for the lighter s quark this is not possible.

If you believe the quark model, and the new quark has charge $+2/3$, like the u quark, then the contribution to the cross-section above the resonances would be $3(2/3)^2R = 4/3\ R$. Then $\sigma(e^-e^+ \rightarrow$ hadrons) above 4 GeV would be $(2 + 4/3)R = 10/3\ R$. The new quark is called 'charm', and written c. Figure 1.9 shows R increases to 5 above 5 GeV, which is greater than 10/3, but there are other processes going on as well as the new quark. In this region, a new third lepton was discovered, called the τ, with its own neutrino [15]. Threshold for τ pair production is 3554 MeV, right in the middle of everything that is going on in Figure 1.9. While technically not a contributor to $\sigma(e^-e^+ \rightarrow$ hadrons), τ pair production (one unit of R well above threshold) leads to complicated final states that were included in the plot until everything got sorted out. The richness of this region was certainly not anticipated, and catapulted e^-e^+ colliders into the forefront of particle physics.

---

[i] Brookhaven and SLAC published simultaneously.

Modern interpretation of the $J/\psi(3095)$ is that it is the ground state of charmonium, the $(c, \bar{c})$ analog of (e^-e^+) bound system positronium. There are intermediate levels between the J/ψ and the ψ' that are fed by the radiative decay of the ψ'. These are the three χ states with $J^{\mathrm{PC}} = 0^{++}, 1^{++}$, and 2^{++}, and masses of 3417, 3510, and 3556 MeV that are fed from $\psi' \rightarrow \chi + \gamma$, followed by $\chi \rightarrow J/\psi + \gamma$. States with the χ quantum numbers cannot be produced directly by e^-e^+ annihilation. The six monoenergetic γ ray energies add in pairs to 589 MeV. A higher vector state, the ψ (3770), decays by strong interaction into $D\bar{D}$, with a D^0 mass of 1864 MeV.

A fourth quark with charge +2/3 was anticipated before the discovery of the J/ψ by an independent argument regarding the nonexistence of 'flavor changing neutral currents', that will be discussed in Chapter 9. Weak transitions like $s \rightarrow d$ do not occur in nature, and can be canceled in the theory if the quarks come in pairs (u, d), and (c, s). Favored weak transitions are $u \rightarrow d$ and $c \rightarrow s$, with the strangeness decays $s \rightarrow u$ suppressed by $\sin(\theta_c)$, where $\theta_c = 13°$ is called the Cabibbo angle [16]. There is a third generation of quarks (t, b) called top and bottom. The b quark has charge $-1/3$, and forms 'bottomonium' in the 10 GeV mass region, with energy level structure similar to charmonium. Bottomonium was discovered in a fixed target experiment at Fermilab similar to the one at Brookhaven that discovered the J, except at higher energy, and measuring $\mu^-\mu^+$

invariant mass rather than e^-e^+ [17]. The $(b, \bar{b})$ bound states are named Υ (upsilon), the lowest energy $\Upsilon(1S)$ has a mass of 9460 MeV. The $\Upsilon(4S)$ at 10,580 MeV decays directly into $(B, \bar{B})$ meson pairs.

As mentioned earlier, there are charmed mesons, called $D^+(c, \bar{d})$, and $D^0(c, \bar{u})$; and bottom mesons $B^+(u, \bar{b})$, and $B^0(d, \bar{b})$, together with their antiparticles. The neutral D's and the neutral B's mix just like the neutral K's, with very interesting consequences.

1.16. Quark Masses

Since there seems to be no free quarks, the quark mass has to be measured indirectly. The numbers in Table 1.6 are constituent quark masses, based on the mass of a three quark baryon. As noted in the table, a spin dependent term is added to the mass formula that splits the pseudoscalar and vector meson masses. Obviously you cannot make a 140 MeV pion out of two 300 MeV quarks without some adjustment. The Gell-Mann–Okubo mass formula Equation (1.34) is written for the meson octets as

- Pseudoscalars

$$\frac{3}{4}M_\eta^2 + \frac{1}{4}M_\pi^2 = M_K^2, \tag{1.59}$$

- Vectors

$$\frac{3}{4}M_\omega^2 + \frac{1}{4}M_\rho^2 = M_{K*}^2. \tag{1.60}$$

Note that the meson masses are squared, and the baryon masses are not. Squared masses seem to work better for the mesons. Also there is an ambiguity regarding what meson replaces the $I = Y = 0$ Λ hyperon, since there are two candidates: η and η' for the pseudoscalars, and ω and ϕ for the vectors. We have already mixed ω and ϕ to get $\phi = (s, \bar{s})$, and using this angle of $-35°$ would increase the ω mass to about 860 MeV, which brings the two sides of Equation (1.60) into better agreement. Equation (1.59) is good to about 5% as written, not as good as the baryon octet, but not bad.

A spin dependent mass term is simpler to apply to the mesons, since there are only two spin 1/2 quarks. Adding a spin dependent term $(\vec{S_1}\cdot\vec{S_2})/(m_1m_2) = (S(S+1)-3/2)/(2m_1m_2)$ gives singlet and triplet formulas:

- Pseudoscalars

$$M_\pi = 2m_u - \frac{3C}{4m_u^2}, \tag{1.61}$$

- Vectors

$$M_\rho = 2m_u + \frac{C}{4m_u^2}. \tag{1.62}$$

Two equations in two unknowns gives $m_u = 306\,\mathrm{MeV}$ and $C/m_u^2 = 630\,\mathrm{MeV}$. Now apply it to a meson with a strange quark:

$$M_K = 306 + m_s - \frac{3}{4} \times 306 \times \frac{630}{m_s}, \tag{1.63}$$

$$M_{K^*} = 306 + m_s + \frac{1}{4} \times 306 \times \frac{630}{m_s}. \tag{1.64}$$

Equation (1.63) is quadratic in m_s with solution $m_s = 486\,\mathrm{MeV}$, and 1.64 then gives $M_{K^*} = 892\,\mathrm{MeV}$, in agreement with experiment. Not bad at all. So this game gives $m_u = m_d = 306\,\mathrm{MeV}$, and $m_s = 486\,\mathrm{MeV}$. This analysis works even better for the baryons. The forms of the baryon magnetic moments are left as an exercise.

The Particle Data Group summarizes all of the data in elementary particle physics every few years [18]. Section 66 concerns the quark masses, written by A. V. Manohar, C. T. Sachrajda, and R. M. Barnett. We quote from their article:

> "Historically the first determinations of quark masses were performed using quark models. [That's us!] The resulting masses only make sense in the limited context of a particular quark model, and cannot be related to quark mass parameters of the Standard Model. In order to discuss quark masses at a fundamental level, definitions based on quantum field theory must be used."

Their article is 15 pages long, and has 64 references to mostly theoretical papers over a ∼20 year time.

Table 1.7: Quark masses.

Quark	Mass (MeV)	Uncertainty (MeV)
u	2.2	$+0.5/-0.4$
d	4.7	$+0.5/-0.3$
s	95	$+9/-3$
c	1275	$+25/-35$
b	4180	$+40/-30$
t	173,000	400

The Particle Data Group then produced the following values as shown in Table 1.7.

These are 'current quark' masses, as opposed to constituent quark masses in Table 1.6. The light quarks are (u, d, s). In order to make a pion mass 140 MeV the light quarks must have substantial kinetic energy. Hadrons are different from any other bound states. A He^3 nucleus weighs less than the sum $2M_p + M_n$. A carbon atom weighs less than $6m_e + M_{C^{12}}$, where the last term is the mass of the carbon-12 nucleus. The mass difference in each case is equal to the binding energy. But a pion weighs more than the masses of the $(q, \bar{q})$ pair. The quarks are confined, and most of the mass is kinetic energy. This phenomenon and its connection to the difference between constituent and current quarks are only partly understood.

The heavy quarks are (c, b, t). Here the difference between current and constituent quarks is smaller. The J/ψ at 3.1 GeV, a bound state of $(c, \bar{c})$, would have constituent quarks of about 1.5 GeV, and Table 1.7 lists a mass of 1.275, not too different. Similarly the mass of the $\Upsilon(1\mathrm{S})$ would imply a b quark mass of about 4.7 GeV, close to the current quark number. So for the heavy quarks the difference is less. The top quark is unique. It is so heavy that it decays into a weak vector boson W and a b quark with a very short lifetime: $t \to W + b$. Color flows from the t to the b (the W boson has no color), and the b hadronizes to make a colorless B meson. The lifetime is so short that toponium $(t, \bar{t})$ cannot form, nor can the top quark form mesons. The top quark appears in the laboratory in the form of its decay products, and its mass can be directly measured. The value quoted in Table 1.7, in contrast with the other quark masses, is a measured quantity.

1.17. Problems

Charge Independence:

1.1 Using the Clebsch–Gordan coefficients for $I = 1/2$ plus $I = 1$, obtain the amplitudes displayed in Equations (1.2)–(1.4).

1.2 The deuteron has I-spin zero, and this leads to several easy tests of charge independence. Show that I-spin conservation predicts:

a. $\sigma(\pi^+ + d \to p + p)/\sigma(\pi^- + d \to n + n) = 1$.
b. $\sigma(p + d \to \mathrm{He}^3 + \pi^0)/\sigma(p + d \to H^3 + \pi^+) = 1/2$ (He^3 and H^3 are an I-spin doublet).
c. $\sigma(d + d \to He^4 + \pi^0) = 0$ (He^4 has I-spin zero).

1.3 Strange particle reactions. The K's are doublets, and the Σ's are triplets, with the Λ a singlet. Obtain the following relations:

a. $\langle \pi^+ p | \mathrm{H}_{3/2} + \mathrm{H}_{1/2} | \Sigma^+ K^+ \rangle = A_{3/2}$.
b. $\langle \pi^- p | \mathrm{H}_{3/2} + \mathrm{H}_{1/2} | \Sigma^- K^+ \rangle = 1/\sqrt{3}(A_{3/2} + 2A_{1/2})$
c. $\langle \pi^- p | \mathrm{H}_{3/2} + \mathrm{H}_{1/2} | \Sigma^0 K^0 \rangle = \sqrt{2}(A_{3/2} - A_{1/2})/\sqrt{3}$.

Identical in form to π p scattering, except the amplitudes are not the same.

d. $\langle \bar{K}^0 p | H_1 + H_0 | \Lambda \pi^+ \rangle = A_1$.
e. $\langle \bar{K}^0 n | H_1 + H_0 | \Lambda \pi^0 \rangle = A_1/\sqrt{2}$.

1.4 G parity. This is a combination of charge conjugation and a rotation of 180° about the y axis in I-spin space: $G = C \times e^{i\pi I_y}$. An even number of pions has even G parity, and an odd number odd. For $I = 1$, the I_y matrix is

$$\begin{pmatrix} 0 & -i/\sqrt{2} & 0 \\ i/\sqrt{2} & 0 & -i/\sqrt{2} \\ 0 & i/\sqrt{2} & 0 \end{pmatrix}.$$

Show that $I_y^3 = I_y$. This does not mean that $I_y^2 = 1$, because I_y is a singular matrix, as are all matrices for integer I-spin. This polynomial can be factored: $I_y(I_y^2 - 1) = 0$, with roots that are the eigenvalues ± 1 and 0. This is a property of all of the I-spin representations. For $I = 1/2$, $I_y^2 - 1/4 = 0$, and for $I = 3/2$, $(I_y^2 - 9/4)(I_y^2 - 1/4) = 0$. All three operators I_x, I_y, and I_z have the same eigenvalues, and obey the same polynomial equation in a given representation. Now show that for $I = 1$:

$$e^{i\theta I_y} = 1 + (\cos\theta - 1)I_y^2 + i\sin\theta I_y. \tag{1.65}$$

Then

$$e^{i\pi I_y} = 1 - 2I_y^2 \tag{1.66}$$

which is the matrix:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix} = \begin{pmatrix} \pi^- \\ -\pi^0 \\ \pi^+ \end{pmatrix} \tag{1.67}$$

Now because $\pi^0 \to \gamma + \gamma$, $C|\pi^0\rangle = +|\pi^0\rangle$, so for consistency $C|\pi^+\rangle = -|\pi^-\rangle$. Then $G|\pi\rangle = -|\pi\rangle$ for all three charge states.

1.5 Evaluate the scattering amplitude $f(\theta)$ — Equation (1.5) — for S and P waves only, i.e. $l = 0$ and $l = 1$. Show that the angular distribution has the form:

$$d\sigma/d\Omega = |f(\theta)|^2 = A + B \times \cos\theta + C \times \cos^2\theta. \tag{1.68}$$

Evaluate the coefficients A, B, and C in terms of δ_0 and δ_1. Now assume $\delta_0 = 0$, and $\delta_1 = \pi/2$, and calculate the total cross-section for relative momentum $\hbar k = 265\,\text{MeV}/c$. Compare your result with the peak value at 1235 MeV in Figure 1.1. The formula 1 $\text{GeV}^{-2} = 0.389\,\text{mb}$ is useful here. There is a small error in that $(2l + 1)$ should be replaced by $(2j + 1)$, or multiply by four instead of three.

Unitary Symmetry:

1.6 Practice with the eight Gell-Mann matrices and I-spin, V-spin, U-spin. Work out some commutation rules:

a. $[I_z, V_\pm] = \pm V_\pm/2$;
b. $[U_+, U_-] = 3Y/2 - I_z \equiv 2U_z$;

c. $[V_+, V_-] = 3Y/2 + I_z \equiv 2V_z$;
d. $[U_+, V_-] = I_-$.

1.7 Check the matrix elements for $\langle \pi^0 | V_- | K^+ \rangle = 1/\sqrt{2}$ and $\langle \pi^0 | U_- | K^0 \rangle = -1/\sqrt{2}$.

1.8 Use color symmetry to show that the baryons around the outside of the octet must have the like quarks in a triplet spin state. Define the baryon magnetic moment:
$\langle \vec{\mu_B} \rangle = \langle S = 1/2, Sz = +1/2 | \sum_q \vec{\mu_q} | S = 1/2, Sz = +1/2 \rangle$, where the sum is over the three quark magnetic moments. The magnetic moment is written as

$$\vec{\mu_q} = g_q \times (q/2m_q) \times \vec{S},$$

where the Lande g factor is defined by this equation, and $q = e_q \times e$.

Obtain the proton spin wave function:

$$p \uparrow = \sqrt{2/3}\, d \downarrow u \uparrow u \uparrow + (d \uparrow (u \uparrow u \downarrow + u \downarrow u \uparrow))/\sqrt{6},$$

in an obvious notation. Calculate the z component of the magnetic moment, and show that $\mu_p = 4/3\mu_u - \mu_d/3$. This formula is true for all octet baryons with two like flavor quarks: $\mu_B = 4/3\mu_2 - 1/3\mu_1$. Use this, plus $\mu_u = -2\mu_d$ (same mass and g factor, charges $e_q = +2/3$ and $-1/3$) to show that $\mu_p/\mu_n = -3/2$. Look up the numbers and compare with experiment. From the values of μ_p and μ_n calculate the quark moments μ_u and μ_d. The baryons in the middle of the octet Σ^0 and Λ have the same quark content (uds). Assume orthogonality, and (ud) in a triplet spin state for Σ^0 matching (uu) and (dd) for Σ^+ and Σ^-. Then the spin and magnetic moment of the Λ comes only from the strange quark. Use the measured value of μ_Λ to obtain μ_s, and then calculate μ_{Σ^+}, μ_{Σ^-}, μ_{Ξ^0}, and μ_{Ξ^-}. Compare results with experiment.

1.9 $\Delta Q = \Delta S$ rule. This rule applies to the weak decays of strange particles. We have seen how $\Delta Q = \Delta S$ forbids the beta decay $\Sigma^+ \rightarrow n + e^+ + \nu_e$ ($\Delta Q = -1, \Delta S = +1$). Apply the $\Delta Q = \Delta S$ rule to the following:

a. $\Xi^0 \rightarrow \Sigma^+ + e^- + \bar{\nu_e}$;
b. $\Xi^0 \rightarrow \Sigma^- + e^+ + \nu_e$;

c. $K^0 \to \pi^+ + e^- + \bar{\nu_e}$;
d. $K^0 \to \pi^- + e^+ + \nu_e$.

1.10 $|\Delta I| = 1/2$ rule. This rule also applies to weak decays of strange particles. For some tests it is useful to use a 'spurion', a particle that has only I-spin 1/2 and nothing else (no energy, spin, etc.). Add in a spurion, and conserve I-spin, and viola! You have the $|\Delta I| = 1/2$ rule. Start with some easy ones. w is the decay rate. Show that the following are consequences of the $|\Delta I| = 1/2$ rule, and compare with experiment:

a. $w(\Lambda \to p + \pi^-)/w(\Lambda \to n + \pi^0) = 2$.
b. $w(K^0 \to \pi^+\pi^-)/w(K^0 \to \pi^0\pi^0) = 2$.
c. Add a spurion to Σ^+ and Σ^- to obtain the amplitudes:
1. $\mathrm{A}(\mathrm{sp} + \Sigma^+ \to p + \pi^0) = \sqrt{2}(A_{3/2} + A_{1/2})/3$;
2. $\mathrm{A}(\mathrm{sp} + \Sigma^+ \to n + \pi^+) = (A_{3/2} - 2A_{1/2})/3$;
3. $\mathrm{A}(\mathrm{sp} + \Sigma^- \to n + \pi^-) = A_{3/2}$.

These are complex numbers that satisfy a 'triangular relation' $\sqrt{2}A_0 + A_+ - A_- = 0$, where the subscript refers to the charge of the pion. Look up the numbers. The magnitudes of the three amplitudes are nearly the same: $|A_0| = |A_+| = |A_-|$, so it is an isoscelese right triangle, with $\sqrt{2}A_0$ the hypotenuse.

1.11 This is a numerical problem. Take Equation (1.26) for the $\bar{K}_0$ content of a beam that starts out pure K_0. Measure time in the K_0 rest frame in units of the K_s lifetime $\tau_s = 0.89510^{-10}$ s. Show that the mass difference term $\Delta mt/\tau_s = 0.47t/\tau_s$, and for $t/\tau_s = 0$ to 10, $t/\tau_L \approx 0$, so that we may write:

$$|\langle\bar{K}_0|K_0\rangle|^2 = (e^{-t} + 1 - 2\cos(0.47t)e^{-t/2})/4; t \text{ in units of } \tau_s. \tag{1.69}$$

Plot this function for $t/s = 0$ to 10. It flattens out at 1/4 for large times. Why?

1.12 Kinematics. Relativity is covered in Chapter 2, but we have already seen some relativistic formulas: s, t, and u. These quantities are invariants, meaning that they are the same in any rest frame. Figure 1.11 shows two typical experimental configurations for studying $2 \to 2$ reactions. Colliding beams

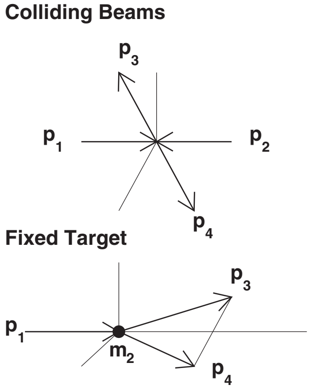


Figure 1.11: Kinematic diagrams for colliding beams and fixed target. In a single ring particle–antiparticle collider the total momentum is zero, and the energy available is twice the ring energy: $\sqrt{s} = 2E$. For a fixed target arrangement $\vec{p_2} = 0$, and $s = m_1^2 + m_2^2 + 2m_2E_1$. The two final state momenta share a plane with the incident beam direction.

are above, and fixed target is below. In a single ring collider, with counter rotating particle and antiparticle beams, the two energies must be equal, $E_1 = E_2$, and the total momentum must be zero, $\vec{p_2} = -\vec{p_1}$, so $s = 4E^2$. Now calculate s for the fixed target arrangement, where $E_2 = m_2$, and $\vec{p_2} = 0$. Show that in the high energy limit, where all masses can be ignored, $s = 2E_1m_2$. The energy available in the collision is $\sqrt{s}$, and $\sqrt{s} = 2E$ for a collider, while $\sqrt{s} = \sqrt{2E_1m_2}$ for fixed target. That is one reason why colliders are so popular.

Calculate the momentum transfer $t = (p_1 - p_3)^2$. Show that if the masses are neglected, $t = -4E^2 \sin^2 \theta/2$. Hence $1/t^2 = 1/(16E^4 \sin^4 \theta/2)$, which is the angle dependence of Rutherford scattering.

1.13 Show that $s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2$.

1.14 Use the formula for s for fixed target to calculate the threshold energy for incident pions in $\pi^- + p \rightarrow \Lambda + K^0$. The threshold is defined as the kinetic energy that must be supplied to make a reaction proceed. At threshold $s = (m_3 + m_4)^2$. $m_3 + m_4$ is the minimum energy in the center of mass, but remember that s is an invariant, and can be calculated in any frame.

References

[1] P. A. M. Dirac, *Proc. Roy. Soc. A* **117**, 778 (1928).

[2] H. Yukawa, *Proc. Phys. Math. Soc. Japan* **17**, 48 (1935).

[3] M. Conversi, E. Pancini, and O. Piccioni, *Phys. Rev.* **68**, 232 (1945).

[4] C. M. G. Lattes, H. Muirhead, G. P. S. Occhialini, and C. F. Powell, *Nature* **159**, 4047 (1947).

[5] G. D. Rochester and C. C. Butler, *Nature* **160**, 855 (1947).

[6] V. E. Barnes *et al.*, *Phys Rev Lett* **12**, 204 (1964).

[7] P. A. Zyla *et al.* (Particle Data Group), *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020); also online at pdg.lbl.gov.

[8] W. Chinowsky and J. Steinberger, *Phys. Rev.* **95**, 1561 (1954).

[9] L. D. Landau, *Doklady Akademii Nauk SSSR* **60**, 207 (1948); see also C. N. Yang, *Phys. Rev.* **77**, 242 (1950).

[10] E. Abouzaid *et al.* (KTeV Collaboration), *Phys. Rev. Lett.* **100**, 182001 (2008).

[11] M. M. Block *et al.*, *Phys. Rev. Lett.* **3**, 291 (1959).

[12] R. F. Schwitters and K. Strauch, *Ann. Rev. Nucl. Sci.* **26**, 118 (1976).

[13] A. M. Boyarski *et al.*, *Phys. Rev. Lett.* **34**, 1357 (1975).

[14] J. J. Aubert *et al.*, *Phys. Rev. Lett.* **33**, 1404; J. E. Augustin *et al.*, *Phys. Rev. Lett.* **33**, 1406 (1974).

[15] M. L. Perl *et al.*, *Phys. Rev. Lett.* **35**, 1489 (1975).

[16] N. Cabibbo, *Phys. Rev. Lett.* **10**, 531 (1963).

[17] S. W. Herb *et al.*, *Phys. Rev. Lett.* **35**, 1489 (1975); W. R. Innes *et al.*, *Phys. Rev. Lett.* **39**, 1240 (1977).

[18] M. Tanabashi *et al.*, *Phys. Rev. D.* **98**, 030001 (2018).

Chapter 2

Special Relativity

2.1. Rotations

We begin our discussion of special relativity, a subject that is a vital part of elementary particle physics, with three-dimensional rotations. Angular momentum operators are the generators of the rotation group, and that group is a subgroup of the Lorentz group that includes Lorentz transformations. The rotation group is closed, because a rotation through 360° gets you back to where you started, but Lorentz transformations are open ended — the Lorentz factor $\gamma = 1/\sqrt{1 - v^2}$ can be any number from one to infinity.

Rotations are familiar in ordinary three-dimensional space. There are various ways to define the three Euler angles (α, β, γ). We adopt the following rotations, according to the right-hand rule: for the thumb in the direction of the rotation axis, the right-hand fingers wrap around the axis in the positive sense:

- Rotate about the z axis through the angle $0 \leq \alpha \leq 360°$.
- Rotate about the y' axis through the angle $0 \leq \beta \leq 180°$.
- Rotate about the z'' axis through the angle $0 \leq \gamma \leq 360°$.

These operations are shown in Figure 2.1. The first rotation leaves the z axis alone; $z' = z$; but the x and y axes are rotated through

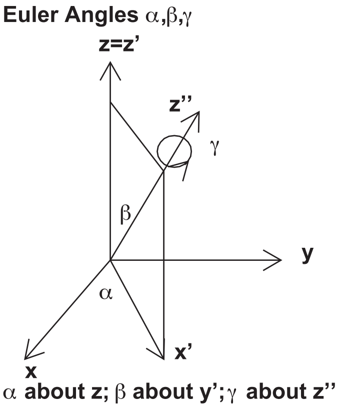


Figure 2.1: Definitions of the three Euler angles described in the text. For a two-body decay, only $\alpha = \phi$ and $\beta = \theta$ need be selected, where ϕ and θ are the usual azimuthal and polar angles of spherical coordinates. For a three-body decay the orientation of one vector is α, β, and the plane of the other two vectors is oriented by γ.

the angle α:

$$x' = x\cos\alpha + y\sin\alpha, \tag{2.1}$$

$$y' = y\cos\alpha - x\sin\alpha, \tag{2.2}$$

$$z' = z. \tag{2.3}$$

The second rotation is in the (z', x') plane through the angle β

$$x'' = x'\cos\beta - z'\sin\beta, \tag{2.4}$$

$$y'' = y', \tag{2.5}$$

$$z'' = z'\cos\beta + x'\sin\beta. \tag{2.6}$$

In terms of the usual spherical coordinates (θ, ϕ), rotate first through $\phi = \alpha$, and then through $\theta = \beta$. The final rotation is about the z'' axis through an angle γ, that changes the orientation of x'' and y''.

$$x''' = x''\cos\gamma + y''\sin\gamma, \tag{2.7}$$

$$y''' = y''\cos\gamma - x''\sin\gamma, \tag{2.8}$$

$$z''' = z''. \tag{2.9}$$

Matrices for the Cartesian coordinates can be multiplied together:

$$\begin{pmatrix} x''' \\ y''' \\ z''' \end{pmatrix} = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \times \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{2.10}$$

The operators that generate these rotations for the spherical harmonics are the components of orbital angular momentum:

$$[L_x, L_y] = L_x L_y - L_y L_x = i\hbar L_z;\ L^2 = L_x^2 + L_y^2 + L_z^2; \quad [L^2, L_x] = [L^2, L_y] = [L^2, L_z] = 0. \tag{2.11}$$

The rotations as defined above are given by the product:

$$R(\alpha, \beta, \gamma) = e^{-i\gamma L_{z''}} \times e^{-i\beta L_{y'}} \times e^{-i\alpha L_z}. \tag{2.12}$$

We have set the factor $\hbar = 1$. The inverse operator has the reverse order, and the opposite signs for the angles:

$$R^{-1}(\alpha, \beta, \gamma) = R(-\gamma, -\beta, -\alpha). \tag{2.13}$$

The sequential axes of rotation that we have adopted help to visualize the rotations, but to calculate matrix elements it is useful to have all of the operators defined in the same set of coordinates. Fortunately, the transformation is simple. You can erase the primes if you invert the order:

$$R(\alpha, \beta, \gamma) = e^{-i\alpha L_z} \times e^{-i\beta L_y} \times e^{-i\gamma L_z}. \tag{2.14}$$

This works because $e^{-i\beta L_{y'}} = e^{-i\alpha L_z} \times e^{-i\beta L_y} \times e^{+i\alpha L_z}$. The rotation matrix elements $\langle l, m'|R(\alpha, \beta, \gamma)|l, m\rangle$ are tabulated like the Clebsch–Gordan coefficients. For $l = 1$ they can be found using the algebra introduced in Problem 1.4.

We do not write two sequential rotations as $e^{-i\alpha L_z - i\beta L_y}$, because that form is incorrect. The usual rule for combining products of exponentials: $e^A e^B = e^{A+B}$, only works if $[A, B] = 0$. The general problem is complicated. It may be studied by defining $f(x) = e^{xA} e^{xB}$, and looking at:

$$df/dx = e^{xA}(A + B)e^{xB} = f(x)(e^{-xB}Ae^{xB} + B). \tag{2.15}$$

Then a series expansion of the A term gives an expression that involves nested commutators:

$$e^{-xB}Ae^{xB} = A + x[A, B] - x^2/(2!)[B, [A, B]] + \cdots \tag{2.16}$$

Integration then gives:

$$\ln f = (A + B)x + x^2/2[A, B] - x^3/6[B, [A, B]] + \cdots \tag{2.17}$$

and in the end set $x = 1$. The nested commutators form an infinite series. The angular momentum operators satisfy $[J_x, J_y] = iJ_z$, writing J instead of L to include 1/2 integer values, so the operators and the commutator do not commute with each other, meaning that the series does not terminate after the correction term $x^2/2[A, B]$. Fortunately, in the case of angular momentum the eigenvalue equations for each representation, mentioned in Chapter 1, come to the rescue, at least for small values of j. Thus, for $j = 1/2$ $J_z^2 - 1/4 = 0$, for $j = 1$ $J_z(J_z^2 - 1) = 0$, and for $j = 3/2$ $(J_z^2 - 1/4)(J_z^2 - 9/4) = 0$. These equations terminate the series expansion of $e^{-i\alpha J_z}$, and all three components satisfy the same polynomial equations. There are important applications where $[A, B] =$ constant. For example, $[x, p] = i\hbar$, and the ladder operators for the harmonic oscillator satisfy $\left[a, a^\dagger\right] = 1$. The nested commutators vanish for these operators, and we may write the simple rule:

$$e^A e^B = e^{(A+B+\frac{1}{2}[A,B])}. \tag{2.18}$$

2.2. Lorentz Transformations

Natural units introduced in Chapter 1 set the velocity of light $c = 1$. Velocity v is now dimensionless, measured in units of c, so that $0 \leq v \leq 1$, with $v = 1$ reserved for massless particles like photons. Similarly the time t is really a length ct, the distance that light travels in that time. We choose the z axis for the direction of the basic Lorentz transformation. The moving (x', y', z') frame travels in the $+z$ direction with velocity v, and the origins of the two three-dimensional frames coincide at $t = 0$.

$$z' = \gamma(z - vt), \tag{2.19}$$

$$t' = \gamma(t - vz), \tag{2.20}$$

$$x' = x, \tag{2.21}$$

$$y' = y, \tag{2.22}$$

$$\gamma = 1/\sqrt{1 - v^2}, \tag{2.23}$$

$$t^2 - z^2 = t'^2 - z'^2. \tag{2.24}$$

The inverse Lorentz transformation has the sign of v reversed: $v \to -v$. Assigning the time to the first component of space–time, we have the following transformation matrix:

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -v\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -v\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}. \tag{2.25}$$

The Lorentz transformation matrix is real and symmetric:

$$L^{\alpha}_{\beta} = L^{*\alpha}_{\beta} = L^{\beta}_{\alpha}, \tag{2.26}$$

$$\det|L| = \epsilon_{\alpha\beta\gamma\delta} \times L^{\alpha}_{0} L^{\beta}_{1} L^{\gamma}_{2} L^{\delta}_{3} = 1, \tag{2.27}$$

$$\epsilon_{\alpha\beta\gamma\delta} = -\epsilon^{\alpha\beta\gamma\delta} = 1, \;\; \alpha, \beta, \gamma, \delta = 0, 1, 2, 3,$$

$$\epsilon_{\beta\alpha\gamma\delta} = -1 \;\; \text{(any two indices interchanged)}, \tag{2.28}$$

$$\epsilon = 0, \;\text{any two indices the same}, \tag{2.29}$$

$$-\epsilon^{\alpha\beta\mu\nu}\epsilon_{\alpha\beta\sigma\tau} = 2(\delta^{\mu}_{\sigma}\delta^{\nu}_{\tau} - \delta^{\mu}_{\tau}\delta^{\nu}_{\sigma}), \tag{2.30}$$

$$-\epsilon^{\alpha\beta\gamma\nu}\epsilon_{\alpha\beta\gamma\tau} = 6\delta^{\nu}_{\tau}, \tag{2.31}$$

$$-\epsilon^{\alpha\beta\gamma\delta}\epsilon_{\alpha\beta\gamma\delta} = 4! = 24, \tag{2.32}$$

$$\delta^{\mu}_{\nu} = 1, \mu = \nu, = 0 \text{ otherwise}. \tag{2.33}$$

We have introduced the totally antisymmetric four-dimensional ϵ tensor. Note that $\epsilon = +1$ for lower indices in cyclic order. This choice agrees with Halzen and Martin [1], and with Bjorken and Drell [2], but Peskin and Schroeder [3] and Barger and Phillips [4] both adopt the opposite convention, i.e. $\epsilon = +1$ for upper indices in cyclic order. Sometimes this convention is important, so be forewarned. The metric tensor introduced below raises and lowers indices in four dimensions. Note that cyclic permutation 0123 → 3012 changes the sign with four indices, but does not change the sign with three, i.e. $\epsilon_{123} = \epsilon_{312}$. The three-dimensional ϵ has no upper indices. A three-dimensional reduction formula is

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}. \tag{2.34}$$

We use Greek letters to specify four-dimensional space–time: $0 \leq \mu \leq 3$, with $\mu = 0$ being the time coordinate. Space coordinates alone are labeled by Roman letters $1 \leq j \leq 3$. Repeated indices are summed over unless otherwise specified. These identities will be useful later.

Coordinate four vectors are written:

$$x^\mu = (x^0, \vec{x}), \tag{2.35}$$

$$x_\mu = (x^0, -\vec{x}), \tag{2.36}$$

$$ds^2 = dx \cdot dx = dx_\mu dx^\mu = g_{\mu\nu} dx^\mu dx^\nu = (dx^0)^2 - d\vec{x}^{\,2}, \tag{2.37}$$

where we have introduced the metric tensor to raise and lower indices:

$$g^{\mu\nu} = g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.38}$$

The metric tensor is its own inverse: $g^2 = 1$. This metric choice is called the Bjorken–Drell metric, after the authors of the textbook *Relativistic Quantum Mechanics* [2]. It maintains real four-vector components at the price of a metric tensor and upper and lower indices. An alternate metric often used has imaginary time as the fourth component $x_\mu = \{\vec{x}, it\}$. There are no upper index vectors. This choice gives time-like vectors with negative length squared: $\vec{x}^2 - t^2$, and leads to factors of $i = \sqrt{-1}$ in various places in a calculation. Of course any physical observable must be metric independent, but the formulas leading up to the final answer can look very different. Emotions can run high regarding the choice of metric. Sakurai in his classic text *Advanced Quantum Mechanics* [5] states that the use of a metric that requires a metric tensor for special relativity, where it is not really necessary, is deplorable. Martinus Veltman in *Diagrammatica* [6] makes a similar clarion call for the complex metric. There are also authors who use a metric tensor, but with the opposite signs, so that the four-vector components are all real, but $ds^2 < 0$ for time like vectors. Despite these headwinds, we proceed.

In our metric, the lower index coordinates Lorentz transform with the inverse velocity $-v$. This is obvious from Equations (2.15) and (2.16). Changing the signs of z and v gives the correct form. Perhaps a more elegant demonstration is

$$gLg = L^{-1}. \tag{2.39}$$

A corollary is that the metric tensor is Lorentz invariant: $LgL = g$ (L is a symmetric matrix).

The gradient operator $\partial/\partial x^{\mu}$ transforms like a lower index vector. This follows from the rules of partial differentiation:

$$\partial/\partial x'^{\mu} = \partial x^{\nu}/\partial x'^{\mu} \times \partial/\partial x^{\nu}, \tag{2.40}$$

$$x^{\nu} = (L^{\nu}_{\mu})^{-1} x'^{\mu}, \tag{2.41}$$

$$\partial x^{\nu}/\partial x'^{\mu} = (L^{\nu}_{\mu})^{-1}, \tag{2.42}$$

$$\partial/\partial x^{\mu} = (\partial/\partial t, \vec{\nabla}). \tag{2.43}$$

The last line shows that the passage to quantum mechanics gives the correct signs:

$$i\partial/\partial x^{\mu} = (i\partial/\partial t, i\vec{\nabla}) = (E, -\vec{p}). \tag{2.44}$$

The minus sign appears on the space momentum vector because it is the lower index four-vector.

This Lorentz transformation can be applied along an arbitrary direction at angles (α, β) relative to the z axis by combining it with rotation matrices:

$$L(\alpha, \beta, v) = R^{-1} \times L(z, v) \times R. \tag{2.45}$$

So, our definition is quite general. The rotation matrices are 3×3, and have no effect on the time component: $R^0_0 = 1, R^0_j = R^j_0 = 0$.

The four-dimensional volume element is Lorentz invariant:

$$d^4x = dx^0 dx^1 dx^2 dx^3 = d^4x'. \tag{2.46}$$

The two volume elements are related by the Jacobian $\det|\partial x'^\mu/\partial x^\nu| = \det|L| = 1$.

2.3. Momentum and Energy

The four momentum of a particle is defined by

$$p^\mu = (E, \vec{p}), \tag{2.47}$$

$$p_\mu = g_{\mu\nu}p^\nu = (E, -\vec{p}), \tag{2.48}$$

$$p \cdot p = p_\mu p^\mu = E^2 - \vec{p}^{\,2} = m^2. \tag{2.49}$$

One advantage of our metric is that square of the length of the momentum four-vector is $+m^2$ rather than $-m^2$. If the velocity of the particle is $v < 1$, then:

$$\gamma = 1/\sqrt{1 - v^2}, \tag{2.50}$$

$$E = \gamma m, \tag{2.51}$$

$$\vec{p} = \gamma m \vec{v}, \tag{2.52}$$

$$\vec{v} = \vec{p}/E. \tag{2.53}$$

If $m = 0$, a case that makes no sense in non-relativistic mechanics, then $v = 1$, and

$$|\vec{p}| = E. \tag{2.54}$$

This relation approximately holds for any massive particle with $E \gg m$. Since the electron mass is only 0.511 MeV, $\gamma = 100$ for a 50 MeV electron, and $|\vec{p}| = E$ to very good accuracy. The four-momentum vector $p^\mu = (E, \vec{p})$ transforms like the coordinates. The Lorentz transformation:

$$p'_z = \gamma(p_z - vE), \tag{2.55}$$

$$E' = \gamma(E - vp_z), \tag{2.56}$$

is appropriate for going to the rest frame of a particle with mass $m > 0$ that is moving with velocity v along the $+z$ direction. In the rest frame $p'_z = 0$, or $v = p_z/E$, which we already knew. Then the second equation gives $E' = m$; the total energy in the particle rest frame is its mass. The particle rest frame is an important concept. A clock sitting at the origin in this frame keeps 'proper time'. Since $z = 0$, Equation (2.20) for proper time τ reads:

$$dt' = \gamma d\tau. \tag{2.57}$$

So, if you are in your rest frame, and the particle is moving along the z axis with velocity v, then a time interval Δt on your clock looks longer than $\Delta\tau$ by the Lorentz factor $\gamma > 1$. This has the important practical consequence that moving radioactive particles live longer in the laboratory frame than in their rest frame. The muon lifetime is 2.2×10^{-6} s. The muon mass $m = 105.6$ MeV. A 10 GeV muon has $\gamma = 100$, and lives 2.2×10^{-4} s in the laboratory. $1/e$ of these muons would travel 66 km before decaying. This phenomenon, called time dilation, allows energetic muons produced high in the atmosphere by cosmic rays to reach the ground. Positive muons from cosmic rays can be stopped in a tank of water, and the time difference between the muon stop and the emission of the positron from the decay $\mu^+ \to e^+ + \nu_e + \bar{\nu_\mu}$ can be measured to obtain the 2.2×10^{-6} s lifetime. It is a nice experiment in physics lab. Unlike people, a senile muon lives just as long as one freshly made.

Length contraction is a related phenomenon. If the velocity of light is constant, and clocks run at different rates, then distances must also change. Let us do a thought experiment firing muons to the moon. The moon is 240,000 miles away, so a light signal takes 1.3 s to make a one way trip. For a muon to make it, we need $\gamma \sim 10^6$, or an energy of 100 TeV, which is admittedly a bit high, but this is a thought experiment. The muon lifetime in the earth rest frame is 2.2 s, so most of them should get there. Now hop on the muon rest frame. The moon, a rather large object, is coming at you at the speed of light! In the muon frame it hits the moon in 1.3×10^{-6} s, and in that time the moon has gone about 0.24 miles. That is length contraction. We use a massive particle like a muon that has $v < c$ by a very small amount. A photon has no rest frame.

2.4. Relativistic Kinematics

Kinematics refers to the conservation of energy and momentum. Dynamics refers to calculations of probabilities, rates, and cross-sections, using quantum mechanics, and is the subject of later chapters. But quite a lot can be learned from kinematics alone, and the subject is very rich and far from trivial. First, consider the two-body fixed target arrangement introduced in Chapter 1. There we learned that the square of the total energy available in the center of mass is

$$s = m_1^2 + m_2^2 + 2m_2E_1. \tag{2.58}$$

The center of mass frame is moving in the laboratory in the direction of the incident beam with velocity $v_c = p_1/(E_1+m_2)$, thus slower than the incident particle with velocity $v = p_1/E_1$. The Lorentz factor for the center of mass is

$$\gamma_c = 1/\sqrt{1 - v_c^2} = (E_1 + m_2)/\sqrt{s} \tag{2.59}$$

or

$$E_1 + m_2 = \gamma_c \times \sqrt{s}. \tag{2.60}$$

The left-hand side is the total energy in the laboratory, and $\sqrt{s}$ is the total energy available in the center of mass, which is smaller by the factor $1/\gamma_c$. It is usually simpler to analyze the two-body to two-body reaction in the center of mass, and transform the final state vectors into the laboratory using (v_c, γ_c) along the beam direction. For $m_1 + m_2 \to m_3 + m_4$, the energy of particle 3 in the center of mass is

$$E_3^c = (s + m_3^2 - m_4^2)/(2\sqrt{s}). \tag{2.61}$$

Energy for particle 4 has m_3 and m_4 interchanged. The two momenta are equal and opposite in the center of mass: $\vec{p_3^c} = -\vec{p_4^c}$. Let $\vec{p_3^c}$ be at polar angle θ_c and azimuthal angle ϕ_c. With the incident beam along the z axis, the azimuthal angle in the (x, y) plane is Lorentz invariant. The polar angle transforms into the laboratory thus,

$$\tan\theta = \sin\theta_c/(\gamma_c(\cos\theta_c + a)), \tag{2.62}$$

where $a = v_c E_3^c / p_3^c$, the ratio of the velocity of the center of mass to the velocity of particle 3 in the center of mass. If $a > 1$, particle 3 never appears in the backward hemisphere in the lab, because the denominator never changes sign. If $a = 1$, particle 3 can just reach 90° in the lab. For $a < 1$, particle 3 can appear in the backward hemisphere; $\cos\theta_c = -a$ maps into $\theta = 90°$. Except for the factor of γ_c in the denominator this formula for $\tan\theta$ is the same as the one in classical mechanics for swimmers in the current of a stream — simple velocity vector addition.

The particle rest frame is appropriate for the analysis of unstable particle decays. In this case, v, γ replace v_c, γ_c in the Lorentz transformation, and $\sqrt{s} = m$, the particle mass. For a two-body decay like $K^+ \to \pi^+\pi^0$ or $K^+ \to \mu^+\nu_\mu$, Equation (2.61) above gives the energy of either final state particle $m \to m_1 + m_2$:

$$E_1 = (m^2 + m_1^2 - m_2^2)/(2m). \tag{2.63}$$

If the masses are equal, then $E_1 = E_2 = m/2$. If $m_2 = 0$, $E_1 = (m^2 + m_1^2)/(2m)$, and $E_2 = (m^2 - m_1^2)/(2m)$. There are no free parameters except for the angles of $\vec{p_1} = -\vec{p_2}$ in the parent rest frame. To generate events for two-body decay in flight, solve for the rest frame momentum and energies, pick the azimuthal angle $0 < \phi < 2\pi$ and polar angle $-1 < \cos\theta < +1$ according to random numbers, and Lorentz transform into the laboratory. Equation (2.62) still holds for the polar angle transformation. The velocity of either final state particle in the parent rest frame is fixed by the kinematics, but v_c, in this case the particle velocity in the laboratory, can be anything $0 < v_c < 1$. The momentum of either particle in the parent rest frame is given by

$$p_1^2 = \lambda(m^2, m_1^2, m_2^2)/(4m^2), \tag{2.64}$$

where $\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2yz - 2xz$ occurs frequently in kinematic expressions. The velocities that enter into the angle transformation are $v_1 = p_1/E_1$, and $v_2 = p_1/E_2$. Of course if $m_2 = 0$, $v_2 = 1$, and photons with energy shifted lower by the Lorentz transformation can appear at 180° in the laboratory frame.

Three-body decays are more complicated to analyze. There are many examples: $K_L^0 \to \pi^+\pi^-\pi^0$, $K^+ \to \pi^0\mu^+\nu_\mu$, $K^+ \to \pi^0 e^+\nu_e$, and all of the hyperon beta decays. In the parent rest frame, there are four equations conserving energy and momentum:

$$\vec{p_1} + \vec{p_2} + \vec{p_3} = 0, \tag{2.65}$$

$$E_1 + E_2 + E_3 = m. \tag{2.66}$$

A little algebra to eliminate particle 3 gives the following equation for the opening angle between particles 1 and 2 in the parent rest frame:

$$\begin{aligned} \cos\theta_{12} = (m^2 + m_1^2 + m_2^2 - m_3^2 + 2E_1E_2 \\ - 2m(E_1 + E_2))/(2p_1p_2). \end{aligned} \tag{2.67}$$

There are two independent variables, E_1 and E_2 that can be picked at random between $0 \le (E_1 - m_1) \le (E_1^{\max} - m_1)$, and similarly for particle 2. The maximum energy for particle 1 comes — provided that $m_1 > 0$, $m_2 > 0$ and $m_3 > 0$ — when particles 2 and 3 'stick together' to form a two-body decay $m \to m_1 + (m_2m_3)$, and the equation for a two-body decay can be used:

$$E_1^{\max} = (m^2 + m_1^2 - (m_2 + m_3)^2)/(2m). \tag{2.68}$$

The equation for $E_2^{\max}$ has the same form with 1 and 2 interchanged. Particles 1 and 2 cannot simultaneously be at the maximum energy because if particle 2 is stuck to particle 3, it is not at its max. This changes if $m_3 = 0$. In that case particles 1 and 2 are simultaneously at their maximum energies if $E_3 = 0$. Particle 3 then plays no role in the energy balance, and the two-body formula can again be applied. Regardless of the values of the masses, the origin on an $(E_1 - m_1, E_2 - m_2)$ plot is excluded. Two particles cannot be simultaneously at rest and conserve momentum. The three particles are collinear on the kinematic boundary. Equation (2.67) gives the formula for the boundary. Set $|\cos\theta_{12}| = 1$, and square both sides to get a quadratic equation in E_2 given E_1. The two solutions give two values for E_2 on the closed boundary curve. Figures 2.2 and 2.3 show the boundary curves for $K \to \mu\pi\nu$ and $K \to e\pi\nu$, respectively.

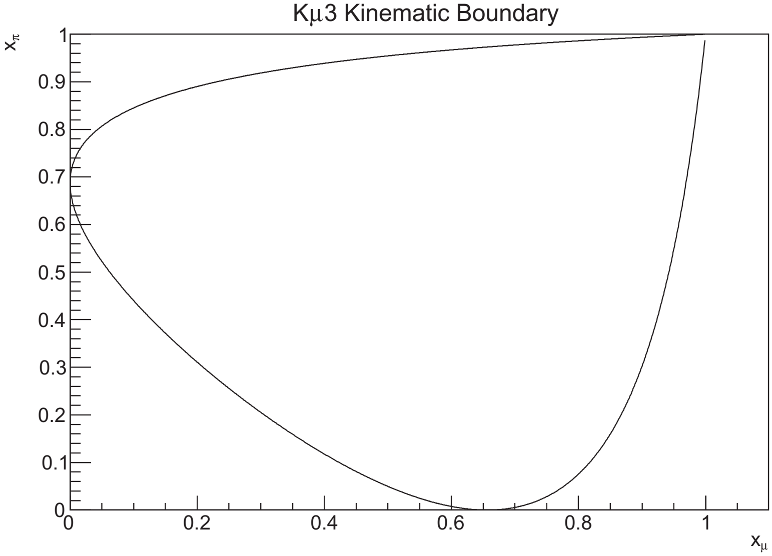


Figure 2.2: Kinematic boundary for $K \to \mu\pi\nu$. The x axis is the muon kinetic energy divided by its maximum value $x_\mu = (E_\mu - m_\mu)/(E_\mu^{\max} - m_\mu)$. The y axis is the same variable for the pion. Because the third particle is massless, the muon and pion have maximum energy simultaneously. The three particles are collinear on the boundary.

In each case, the neutrino mass is zero. The quadratic equation from Equation (2.67) for $K\mu3$ is

$$\begin{aligned} &AE_\pi^2 + BE_\pi + C = 0, \\ &A = (m_K^2 + m_\mu^2 - 2m_K E_\mu), \\ &B = (m_K - E_\mu)(2m_K E_\mu - M^2), \\ &C = (M^2 - 2m_K E_\mu)^2/4 + m_\pi^2(E_\mu^2 - m_\mu^2), \end{aligned} \tag{2.69}$$

with $M^2 = m_K^2 + m_\mu^2 + m_\pi^2$.

A way to generate events consistent with the kinematics is to pick $(E_1 - m_1)/(E_1^{\max} - m_1)$ and $(E_2 - m_2)/(E_2^{\max} - m_2)$ each at random between zero and one, and substitute the energies into Equation (2.67), which is easy to evaluate with a computer. If $|\cos\theta_{12}| < 1$, keep the event. Otherwise, throw it out and start over. The distribution of dots on the (E_1, E_2) plot will all lie inside the boundary. Each dot represents a possible three-body decay configuration in a plane that is randomly oriented in the parent rest

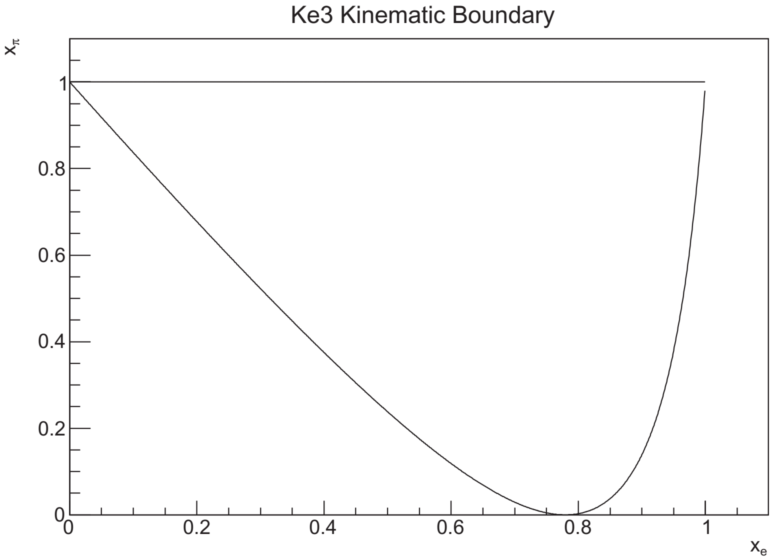


Figure 2.3: Kinematic boundary for $K \to e\pi\nu$. The curve is flat on top, because the electron is treated as massless, like the neutrino, and the two can share energy along the top with the maximum pion kinetic energy.

frame. Pick the Euler angles at random, and Lorentz transform each momentum vector into the laboratory for a moving parent particle.

We will need Dirac's delta function:

$$\delta(x) = \delta(-x), \tag{2.70}$$

$$\delta(x - x_0) = 0, x \neq x_0, \tag{2.71}$$

$$\int_{x_0-\epsilon}^{x_0+\epsilon} \delta(x - x_0)dx = 1, \tag{2.72}$$

$$\int_{x_0-\epsilon}^{x_0+\epsilon} f(x)\delta(x - x_0)dx = f(x_0), \tag{2.73}$$

$$\int_{x_0-\epsilon}^{x_0+\epsilon} f(x)(d\delta(x - x_0)/dx)dx = -df/dx|_{x=x^0}, \tag{2.74}$$

$$\delta(ax) = (1/a)\delta(x), \tag{2.75}$$

$$\delta(x^2 - a^2) = (1/2a)(\delta(x - a) + \delta(x + a)). \tag{2.76}$$

If $\theta(x-x_0)$ is a unit step function: $\theta(x-x_0) = 0, x < x_0$; $\theta(x-x_0) = 1, x > x_0$, then $\delta(x-x_0) = d\theta/dx$. There are many different algebraic forms of the δ function that we will encounter in this book.

We have seen that d^4x is Lorentz invariant. This is true also for d^4p for the same reason. Hence we may write an invariant integral:

$$\int \delta(p^2 - m^2) d^4p. \tag{2.77}$$

Now write the argument of the δ function thus,

$$p^2 - m^2 = E^2 - (\vec{p}^{\,2} + m^2) = E^2 - E_0^2, \tag{2.78}$$

$$\begin{aligned}\int \delta(p^2 - m^2) d^4p &= \int d^3p dE \delta(E^2 - E_0^2)\\ &= \int d^3p \int dE (1/2E_0)(\delta(E - E_0)\\ &\quad + \delta(E + E_0)).\end{aligned} \tag{2.79}$$

In the last step, we have used a property of the δ function listed above. We drop the second term, that corresponds to negative total energy, to arrive at the invariant expression:

$$d^3p/2E \text{ is Lorentz invariant.} \tag{2.80}$$

Golden Rule #2 for the transition rate from time dependent perturbation theory $w_{i\to f}$ in quantum mechanics is

$$w_{i\to f} = (2\pi/\hbar) |\langle f| H' |i\rangle|^2 \times dN/dE. \tag{2.81}$$

Golden Rule #1, in case you are curious, is the second-order perturbation formula with the energy denominator. The matrix element for the perturbation Hamiltonian $\langle f| H' |i\rangle = \int d^3x \psi_f^* H' \psi_i$ has energy units, so the transition rate has units $w = E^2/(E^2 \times t) = 1/t$. The lifetime of the initial state τ is given by the reciprocal of the sum over all possible final states: $1/\tau = \sum_f w_{i\to f}$. The lifetime, as we have seen, is not Lorentz invariant. The last factor in Equation (2.81), dN/dE, is called the density of final states, or phase space. We will introduce a relativistic version of Golden Rule #2 later, but for now we concentrate on the density of final states. Lorentz invariant phase space for a three-body final state is

$$(1/2\pi)^9(d^3p_1/(2E_1))(d^3p_2/(2E_2))(d^3p_3/(2E_3)) \\ \times \delta(\vec{p_1}+\vec{p_2}+\vec{p_3})\delta(E_1+E_2+E_3-m). \qquad (2.82)$$

The δ functions conserve momentum and energy. To eliminate particle 3, integrate $\int d^3p_3/(2E_3)\delta(\vec{p_1}+\vec{p_2}+\vec{p_3}) = 1/(2E_3)$, where $E_3 = \sqrt{(\vec{p_1}+\vec{p_2})^2+m_3^2}$. This leaves us with particles 1 and 2, and an energy δ function.

$$\text{Invariant phase space} = (1/2\pi)^9(d^3p_1/(2E_1))(d^3p_2/(2E_2)) \\ \times \delta(E_1+E_2+E_3-m)/(2E_3). \quad (2.83)$$

The conservation of momentum is contained in

$$E_3 = \sqrt{\vec{p_1}^2+\vec{p_2}^2+2\vec{p_1}\cdot\vec{p_2}+m_3^2}.$$

So E_3 contains the cosine of an angle between the two momentum vectors $\vec{p_1}$ and $\vec{p_2}$, $\cos\theta_{12}$. Equation (2.67) comes from solving the argument of the energy δ function for $\cos\theta_{12}$. There are two solid angles left in the problem, $d\Omega_1$ and $d\Omega_2$. We may integrate over one of them, and write the other as $d\Omega_{12}$. After integration over the azimuth we have:

$$\text{Invariant phase space} \\ = 2/(2\pi)^7(p_1^2dp_1/(2E_1))(p_2^2dp_2/(2E_2)) \\ \times (d\cos\theta_{12}/(2E_3)\delta(E_1+E_2+E_3-m). \qquad (2.84)$$

The integral over $\cos\theta_{12}$, which appears in the argument of the δ function, can be treated by changing the variables:

$$\delta(f(x))dx = \delta(f)df \times (dx/df); \quad \int \delta(f)df = 1. \qquad (2.85)$$

In this case, with $x = \cos\theta_{12}, df/dx = |\vec{p_1}||\vec{p_2}|/(E_3)$. Integration over dx then gives the desired result, after using $p_1dp_1 = E_1dE_1$:

$$\text{Invariant three-body phase space} \sim dE_1dE_2. \qquad (2.86)$$

Therefore in the absence of any forces on the particles that may distort the distribution, the event dots should uniformly populate the (E_1, E_2) plot. Again, a point on the plot completely specifies the three-body decay kinematics, apart from the orientation of the decay plane in the parent center of mass. The three Euler angles (α, β, γ) introduced at the beginning of this chapter are actually useful in this regard. Pick one momentum vector, $\vec{p_1}$ for instance, oriented at angles α and β relative to the fixed axes in the rest frame (z-axis in the direction of the motion, and (x, y) in a plane perpendicular to z), then orient the $\vec{p_2}, \vec{p_3}$ plane with the angle γ. This gives the components (p_x, p_y, p_z) for each particle in the rest frame. Lorentz transform the p_z's into the laboratory, leaving p_x and p_y invariant. There is more discussion of phase space techniques in the Appendix.

2.5. Relativistic Wave Equations

Erwin Schrödinger's famous wave equation needs no apology. It stands right next to Maxwell's equations and Newton's laws as one of the most successful equations in physics. For one particle, in cgs units, it reads:

$$-\hbar^2/(2m)\nabla^2\psi(\vec{x}, t) + V(\vec{x})\psi(\vec{x}, t) = i\hbar\partial\psi(\vec{x}, t)/\partial t. \tag{2.87}$$

It follows from the non-relativistic energy relation: $p^2/(2m)+V(\vec{x}) = E$, with the operator substitution $\vec{p} \to -i\hbar\vec{\nabla}$ and $E \to i\hbar\partial/\partial t$, and the introduction of a complex wave function $\psi(\vec{x}, t)$. The probability density $\rho = \psi^*\psi$, and the associated probability current density $\vec{j} = \hbar/(2mi)\times(\psi^*\vec{\nabla}\psi - (\vec{\nabla}\psi^*)\psi)$ satisfy a continuity equation that follows from the wave equation, provided that the potential is real:

$$\vec{\nabla} \cdot \vec{j} + \partial\rho/\partial t = 0. \tag{2.88}$$

If the wave function ψ is real, which it can be for bound states, then $\vec{j} = 0$, and $\rho = \psi^2$ is constant in time. This is appropriate for a bound state, in which the particle is not going anywhere. The only shortfall of Equation (2.87) is that it cannot be relativistically invariant. The Hamiltonian operator $H = -\hbar^2/(2m)\vec{\nabla}^2 + V(\vec{x})$, is an energy, that transforms like the time component of a four-vector. What are the other three components? Potentials $V(\vec{x})$ are problematic in relativity. If they are time components of four vectors, then the other three components have to be included. If they are four-dimensional scalars, then they can be added to the mass. Then there is the difficulty with the operators, quadratic in space but linear in time — an obvious problem for Lorentz transformations that mix space and time coordinates.

Schrödinger knew this, and wrote down a free particle wave equation based on the same operators, but the energy relation $E^2 = \vec{p}^{\,2} + m^2$:

$$(\nabla^2 - \partial^2/\partial t^2 - m^2)\phi(x) = 0. \tag{2.89}$$

In units $\hbar = c = 1$. This equation, named after Oskar Klein and Walter Gordon, has plane wave solutions of the form $\phi(x) = e^{-ip\cdot x} = e^{i(\vec{p}\cdot\vec{x}-Et)}$. Substitution into the wave equation gives $E^2 - \vec{p}^{\,2} - m^2 = 0$. Like the Schrödinger equation, Equation (2.89) contains the conservation of a current. Write down the equation for ϕ^*, multiply by ϕ from the right. Then multiply Equation (2.89) from the left by ϕ^*, and subtract the two. $\vec{\nabla} \cdot \vec{j} + \partial\rho/\partial t = 0$, then follows, with

$$\vec{j} = (\phi^*\vec{\nabla}\phi - (\vec{\nabla}\phi^*)\phi)/(2i), \tag{2.90}$$

$$\rho = ((\partial\phi^*/\partial t)\phi - \phi^*\partial\phi/\partial t)/(2i). \tag{2.91}$$

The common factors of $1/(2i)$ are added for convenience, as the two expressions are imaginary. The form for $\vec{j}$ matches the non-relativistic one when multiplied by $\hbar/m$. The Klein–Gordon equation is real, so it has real solutions, for which both $\vec{j}$ and ρ vanish. The non-relativistic version of the density ρ can be retrieved by first dividing by the mass to match $\vec{j}$, and then factoring out the mass term from the energy by writing: $\phi(\vec{x},t) = \psi(\vec{x},t)e^{-imt}$. After differentiation discard the terms proportional to $1/m$, leaving the non-relativistic result $\rho = \psi^*(\vec{x},t)\psi(\vec{x},t)$. The relativistic expression for ρ is not positive definite, because of the second time derivative in the wave equation. This dilemma is solved by multiplying ρ and $\vec{j}$ by the elementary charge e, and interpreting the relation as conservation of charge. The equation is suitable for spinless particles like a pair of charged pions, π^+ and π^-. The wave function $\phi(x)$ must be complex to describe a charged pair, otherwise $\rho = 0$ as we have seen. The equation can be written in covariant form by introducing the four-dimensional Laplacian:

$$g^{\mu\nu}\partial/\partial x^\mu \partial/\partial x^\nu = \Box^2 = \partial^2/\partial t^2 - \nabla^2, \tag{2.92}$$

$$(\Box^2 + m^2)\phi(x) = 0. \tag{2.93}$$

2.6. Dirac's Equation

Dirac tried a different approach. In order to have an invariant form, and preserve a positive definite probability density, Dirac wrote down an equation linear in momentum and energy, rather than quadratic. The equation has many other benefits, as we shall see. In our units Dirac's original free particle equation is

$$(\vec{\alpha} \cdot \vec{p} + \beta m)\psi(x) = E\psi(x). \tag{2.94}$$

This equation looks deceptively innocent. First require $E^2 = \vec{p}^{\,2} + m^2$ by squaring both sides:

$$(\vec{\alpha} \cdot \vec{p} + \beta m)(\vec{\alpha} \cdot \vec{p} + \beta m) = E^2, \tag{2.95}$$

$$(\alpha_i\alpha_j + \alpha_j\alpha_i)p_i p_j/2 + (\beta\alpha_j + \alpha_j\beta)mp_j + \beta^2 m^2 = \vec{p}^{\,2} + m^2, \tag{2.96}$$

$$(\alpha_i\alpha_j + \alpha_j\alpha_i) = 0,\ i \neq j,\ \alpha_i^2 = 1, \tag{2.97}$$

$$\beta\alpha_j + \alpha_j\beta = 0,\ \beta^2 = 1. \tag{2.98}$$

The last two lines come from comparing right- and left-hand sides of the second equation. If α_i and β are matrices, they are Hermitian: $\vec{\alpha}^\dagger = \vec{\alpha}$, and $\beta^\dagger = \beta$, because the Hamiltonian operator must be Hermitian. They must have an even number of rows and columns, because $\mathrm{Tr}\{\alpha_i\} = \mathrm{Tr}\{\beta\} = 0$. The Pauli spin matrices have these properties. They are Hermitian; they anticommute, are traceless, and have square 1:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.99}$$

Any 2×2 Hermitian matrix can be formed from these three matrices, plus the unit matrix. The unit matrix commutes with all other 2×2

matrices, so there are only three σ's, and we need four. The smallest size that satisfies the conditions is 4×4. Matrices that work are:

$$\alpha_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}; \quad \alpha_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix};$$

$$\alpha_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}; \quad \beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.100}$$

The Dirac wave function is a column vector with four components:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}. \tag{2.101}$$

In a compressed notation these matrices are:

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}; \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.102}$$

The four-component wave function $\psi(x)$ is compressed into two two-component functions ψ_a and ψ_b, and we write the Dirac equation as

$$\begin{pmatrix} m - E & \vec{\sigma} \cdot \vec{p} \\ \vec{\sigma} \cdot \vec{p} & -m - E \end{pmatrix} \begin{pmatrix} \psi_a \\ \psi_b \end{pmatrix} = 0. \tag{2.103}$$

The two-component column vectors ψ_a and ψ_b satisfy the equations:

$$\psi_a = \vec{\sigma} \cdot \vec{p}/(E - m)\psi_b, \tag{2.104}$$

$$\psi_b = \vec{\sigma} \cdot \vec{p}/(E + m)\psi_a. \tag{2.105}$$

The two equations are consistent, because $(\vec{\sigma} \cdot \vec{p})(\vec{\sigma} \cdot \vec{p}) = \vec{p}^{\,2}$. Normally $E > m$ and positive, so ψ_a is called the 'large component', while ψ_b has $2m$ in the denominator, and is small in the non-relativistic limit, i.e. $E < 2m$. The quadratic equation $E^2 = \vec{p}^{\,2} + m^2$ allows for negative energy solutions, a feature to which we will return. If $E < 0$ the roles of large and small components interchange. The presence of the Pauli spin matrices in the wave equation is strong evidence that it applies to spin 1/2 particles, and indeed Dirac's equation describes electrons.

The adjoint wave function is a four component row vector:

$$\psi^\dagger = \begin{pmatrix} \psi_1^* & \psi_2^* & \psi_3^* & \psi_4^* \end{pmatrix}. \tag{2.106}$$

Substituting the operators $\vec{p} \to -i\vec{\nabla}$, and $E \to i\partial/\partial t$ gives the equations:

$$-i\vec{\alpha} \cdot \vec{\nabla}\psi + m\beta\psi = i\partial\psi/\partial t, \tag{2.107}$$

$$i\vec{\nabla}\psi^\dagger \cdot \vec{\alpha} + m\psi^\dagger\beta = -i\partial\psi^\dagger/\partial t. \tag{2.108}$$

Multiply the top equation from the left by $\psi^\dagger$ and the bottom one from the right by ψ and subtract to get the continuity equation:

$$\vec{\nabla} \cdot (\psi^\dagger\vec{\alpha}\psi) + \partial(\psi^\dagger\psi)/\partial t = 0. \tag{2.109}$$

Thus, one of Dirac's goals, a positive definite probability density $\rho = \psi^\dagger\psi$, is achieved. The probability current density $\vec{j} = \psi^\dagger\vec{\alpha}\psi$ looks a bit strange, since $\vec{\alpha}$ is just a matrix with eigenvalues ± 1, apparently unrelated to particle velocity $|\vec{v}| < 1$. The usual quantum mechanical rule for the time dependence of an operator: $dx/dt = -i[x, H] = \alpha_x$ may help in interpretation of this result. We will return to this question below.

The covariant form of the Dirac equation can be obtained by multiplying Equation (2.107) from the left by β:

$$-i\beta\vec{\alpha} \cdot \vec{\nabla}\psi + m\psi = i\beta\partial\psi/\partial t. \tag{2.110}$$

Rearrange terms, define new 4×4 matrices $\vec{\gamma} = \beta\vec{\alpha}$ and $\gamma^0 = \beta$:

$$\vec{\gamma} \cdot \vec{\nabla}\psi + \gamma^0\partial\psi/\partial t = -im\psi, \tag{2.111}$$

$$\gamma^\mu\partial\psi/\partial x^\mu = -im\psi. \tag{2.112}$$

The last line is the covariant form that is inscribed in a stone on the floor of Westminster Abbey in London, near the tomb of Isaac Newton, in memory of Paul Dirac. The matrices γ^μ satisfy the anticommutation rules:

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}. \tag{2.113}$$

The space components are not Hermitian:

$$(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0. \tag{2.114}$$

So, $(\gamma^j)^\dagger = -\gamma^j$, but $(\gamma^0)^\dagger = +\gamma^0$. Also, $(\gamma^j)^2 = -1$, $(\gamma^0)^2 = +1$, and all the γ^μ anticommute. These matrices look like components of a four-vector, but they are not. They are the same matrices in any Lorentz frame. Or to put it another way, a unitary transformation can change the forms of the matrices, independent of the Lorentz frame. The Lorentz transformation properties of the Dirac equation will be discussed in a later section. The facts that there are four components to space–time, and there are four components to the Dirac equation seem to be an accidental coincidence, with no deeper meaning.

Write the Dirac four component wave function as

$$\psi(x) = u(p)e^{-ip\cdot x}. \tag{2.115}$$

Substitution into Dirac's equation $\gamma^\mu\partial\psi/\partial x^\mu = -im\psi$ gives the equation for the spinor $u(p)$:

$$(\gamma^\mu p_\mu - m)u(p) = 0. \tag{2.116}$$

Introduce Feynman's 'slash' notation:

$$\not{p} \equiv \gamma^\mu p_\mu. \tag{2.117}$$

Then the equation for $u(p)$, the plane wave solution to the Dirac equation, reads:

$$(\not{p} - m)u(p) = 0. \tag{2.118}$$

The adjoint equation is

$$u^\dagger(p)(p_\mu\gamma^{\mu\dagger} - m) = 0. \tag{2.119}$$

Multiply from the right by γ^0, and use $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$ plus $(\gamma^0)^2 = 1$ and $\bar{u}(p) = u^\dagger(p)\gamma^0$ to obtain:

$$\bar{u}(p)(\not{p} - m) = 0. \tag{2.120}$$

Hard to imagine fundamental equations simpler than these. For a neutrino Equation (2.118) becomes just: $\not{p}u(p) = 0$. Spinors have four components. We adopt the definition and normalization of Halzen and Martin's classic text *Quarks and Leptons* [1]. For positive energy $E > 0$ and spin up in the non-relativistic limit:

$$u^{(1)}(p) = \sqrt{E+m}\begin{pmatrix} 1 \\ 0 \\ p_z/(E+m) \\ (p_x + ip_y)/(E+m) \end{pmatrix}. \tag{2.121}$$

For spin down:

$$u^{(2)}(p) = \sqrt{E+m}\begin{pmatrix} 0 \\ 1 \\ (p_x - ip_y)/(E+m) \\ -p_z/(E+m) \end{pmatrix}. \tag{2.122}$$

The normalization chosen is

$$u^{(1)\dagger}(p) \times u^{(1)}(p) = u^{(2)\dagger}(p) \times u^{(2)}(p) = 2E, \tag{2.123}$$

and for the adjoint spinors $\bar{u}(p) = u^{\dagger}(p)\gamma^0$ — an operation that changes the sign of the two lower components:

$$\bar{u}^{(1)}(p) \times u^{(1)}(p) = \bar{u}^{(2)}(p) \times u^{(2)}(p) = 2m. \tag{2.124}$$

The cross terms are orthogonal:

$$\bar{u}^{(2)}(p) \times u^{(1)}(p) = u^{(2)\dagger}(p) \times u^{(1)}(p) = 0. \tag{2.125}$$

In covariant form the probability density $\psi^{\dagger}\psi$ transforms as the time component of a four-vector, while $\bar{\psi}\psi$ is a scalar, like the mass. This explains the above choice of spinor normalization. Returning to the interpretation of $\vec{\alpha}$ as a velocity operator, we calculate the expectation value $\langle\psi^{\dagger}\alpha_z\psi\rangle$ for positive energy spinors. The Fourier expansion of $\psi(x)$ can be written as follows:

$$\psi(x) = (1/(2\pi)^3)\int d^3p/2E\sum_{s=1}^{2}(u^{(s)}(p)a^{(s)}(p)e^{i(\vec{p}\cdot\vec{x}-Et)}). \tag{2.126}$$

We have adopted this form, called the continuum normalization, for all Fourier expansions in this book. Some texts use the box normalization instead, where the wave numbers are quantized by periodic boundary conditions on a large cubic box of dimension $V = L^3$, thus $k_x^{n_x} = (2\pi n_x)/L$, n_x being a (large) integer. This condition gives running waves that have the same phase at $x = \pm L/2$, and form a continuous wave for boxes placed end to end. For box normalization, the $\int d^3k$ is replaced by a sum: $\sum_{n_x,n_y,n_z}$, and the delta functions that we will encounter are replaced by δ_{n_x,n'_x} Kronecker δ's. Boundary conditions on infinite walls give standing waves for real wave functions, and give the same answers for the enumeration of the states. Thus, $\sin(kx) = 0, x = \pm L/2$ means $k_x L/2 = n_x \pi$. It is interesting to note that this condition can be used only with second-order differential equations, like the Schrödinger equation. The Dirac equation is first order, and if e^{ikx} is a solution, then e^{-ikx} is not.

In Equation (2.126), $a^{(s)}(p)$ is the Fourier amplitude with momentum $\vec{p}$ and spin s. The calculation of the expectation value of α_z using the above spinors is straightforward.

$$\begin{aligned}\psi^\dagger(x)\alpha_z\psi(x) = (1/(2\pi)^6)&\int d^3p/2E\int d^3p'/2E'\\ &\times\sum_{s=1}^{2}\sum_{s'=1}^{2} a^{(s)}(p)a^{*(s')}(p')u^{\dagger(s')}(p')\alpha_z u^{(s)}(p)\\ &\times e^{i((\vec{p}-\vec{p'})\cdot\vec{x}-(E-E')t)}.\end{aligned} \tag{2.127}$$

The integral over all space $(1/(2\pi)^3)\int d^3x e^{i(\vec{p}-\vec{p'})\cdot\vec{x}} = \delta(\vec{p}-\vec{p'})$ is verified in Chapter 3; then $\int d^3p'/2E'$ replaces $\vec{p'}$ with $\vec{p}$, the time dependence vanishes, and $u^{\dagger(s')}(p)\alpha_z u^{(s)}(p) = 0$ for $s' \neq s$.

$$\begin{aligned}\langle\psi^\dagger\alpha_z\psi\rangle &= \int d^3x\psi^\dagger(x)\alpha_z\psi(x)\\ &= (1/(2\pi)^3)\int d^3p/2E\sum_{s=1}^{2}|a^{(s)}(p)|^2(p_z/E).\end{aligned} \tag{2.128}$$

Hence the expectation value of α_z for positive energy spinors is just the weighted average of $v_z = p_z/E$.

The negative energy spinors are $u^{(3)}(p)$ and $u^{(4)}(p)$, defined in terms of $|E|$:

$$u^{(3)}(p) = \sqrt{|E|+m}\begin{pmatrix} -p_z/(|E|+m) \\ -(p_x+ip_y)/(|E|+m) \\ 1 \\ 0 \end{pmatrix};$$

$$u^{(4)}(p) = \sqrt{|E|+m}\begin{pmatrix} (-p_x+ip_y)/(|E|+m) \\ p_z/(|E|+m) \\ 0 \\ 1 \end{pmatrix}. \qquad (2.129)$$

These spinors are normalized thus:

$$u^{(3)\dagger}(p) \times u^{(3)}(p) = u^{(4)\dagger}(p) \times u^{(4)}(p) = 2|E|, \qquad (2.130)$$

and

$$\bar{u}^{(3)}(p) \times u^{(3)}(p) = \bar{u}^{(4)}(p) \times u^{(4)}(p) = -2m. \qquad (2.131)$$

The cross terms are again orthogonal. Since $E = |E|$ for the positive energy spinors $u^{(1)}$ and $u^{(2)}$, positive and negative energy spinors have the same normalization for $u^\dagger u$. The sign flips for $\bar{u}u$. The expectation value of α_z for negative energy spinors proceeds just like the positive energy above, with the answer having the opposite sign: $\langle\alpha_z\rangle = -\langle p_z/|E|\rangle$. The negative energy electron spinors are identified as positive energy positron spinors at the end of this chapter.

2.7. Dirac's Equation in an Electromagnetic Field

The Dirac Hamiltonian is $H = \vec{\alpha}\cdot\vec{p} + \beta m$. We may include electromagnetic potentials in the same way as done in the Schrödinger equation, namely,

$$\vec{p} \to \vec{p} - e\vec{A}, \tag{2.132}$$

$$E \to E - eV. \tag{2.133}$$

One of Maxwell's equations is $\vec{\nabla}\cdot\vec{B} = 0$, permitting the vector potential $\vec{A}$ to be introduced as a calculational aid, defined by $\vec{B} = \vec{\nabla} \times \vec{A}$. The vector and scalar potentials $(\vec{A}, V)$ are often simpler to use than the fields $(\vec{E}, \vec{B})$ in problems in electricity and magnetism. The Lagrangian and Hamiltonian formulations of mechanics are based on the potentials. A common form of $\vec{A}$ for a constant magnetic field $\vec{B}$ is $(A_x = -By/2, A_y = Bx/2, A_z = 0)$, so $\vec{\nabla} \times \vec{A} = B\hat{z}$. The vector potential is not uniquely determined by its definition. Any arbitrary well-behaved function $\phi(x)$ can be added to $\vec{A'} = \vec{A} + \vec{\nabla}\phi$ without changing the magnetic field. The electric field is similarly invariant with the replacement $V \to V - \partial\phi/\partial t$. This is called a gauge transformation, and any results are required to be gauge invariant. The vector potential $\vec{A}$ and gauge invariance play an important role in quantum electrodynamics. $\vec{A}$ is no longer just a calculational aid. The gradient of the scalar potential, or the voltage, gives the electric field, with a term $\partial\vec{A}/\partial t$ subtracted to satisfy Faraday's law:

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho, \tag{2.134}$$

$$\vec{\nabla} \cdot \vec{B} = 0, \tag{2.135}$$

$$\vec{\nabla} \times \vec{E} = -\partial\vec{B}/\partial t, \tag{2.136}$$

$$\vec{B} = \vec{\nabla} \times \vec{A}, \tag{2.137}$$

$$\vec{E} = -\vec{\nabla}V - \partial\vec{A}/\partial t. \tag{2.138}$$

The Dirac equation in an external electromagnetic field is

$$(\vec{\alpha} \cdot (\vec{p} - e\vec{A}) + \beta m)\psi = (E - eV)\psi \tag{2.139}$$

or

$$(\vec{\alpha} \cdot (\vec{p} - e\vec{A}) + \beta m - E + eV)\psi = 0. \tag{2.140}$$

In order to study this equation in the nonrelativistic limit, it is useful to multiply through from the left with the operator

$\vec{\alpha}\cdot(\vec{p}-e\vec{A})+\beta m+E-eV$. This operation gives several terms:

$$\begin{aligned}[\vec{\alpha}\cdot(\vec{p}-e\vec{A})&\times(\vec{\alpha}\cdot(\vec{p}-e\vec{A}))+(\beta m+(E-eV))\\&\times(\beta m-(E-eV))+(\vec{\alpha}\cdot(\vec{p}-e\vec{A}))\times(\beta m-(E-eV))\\&+(\beta m+(E-eV))\times(\vec{\alpha}\cdot(\vec{p}-e\vec{A}))]\psi=0.\end{aligned}\tag{2.141}$$

The first term can be simplified by using an identity for the α^j matrices similar to the one for the Pauli spin matrices: $(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma})=\vec{a}\cdot\vec{b}+i\vec{\sigma}\cdot(\vec{a}\times\vec{b})$, where $\vec{a}$ and $\vec{b}$ commute with $\vec{\sigma}$, but not necessarily with each other. The α^j analog is

$$(\vec{\alpha}\cdot\vec{a})(\vec{\alpha}\cdot\vec{b})=\vec{a}\cdot\vec{b}+i\vec{\Sigma}\cdot(\vec{a}\times\vec{b}),\tag{2.142}$$

where $\vec{\Sigma}$ is the four-dimensional Pauli spin matrix:

$$\vec{\Sigma}=\begin{pmatrix}\vec{\sigma} & 0\\ 0 & \vec{\sigma}\end{pmatrix}.\tag{2.143}$$

After doing a little algebra and substituting $\vec{p}\to -i\vec{\nabla}$ the first term becomes:

$$\vec{\alpha}\cdot(\vec{p}-e\vec{A})\times(\vec{\alpha}\cdot(\vec{p}-e\vec{A}))=(\vec{p}-e\vec{A})^2-e\vec{\Sigma}\cdot(\vec{\nabla}\times\vec{A}).\tag{2.144}$$

Since $\vec{B}=\vec{\nabla}\times\vec{A}$, the last term has the form of the energy of an intrinsic magnetic moment with the external magnetic field. The other terms in Equation (2.141) are straightforward, keeping in mind that many operators do not commute, and that both $\vec{A}$ and V are functions of $\vec{x}$ and t. Substitute $\vec{p}=-i\vec{\nabla}$, and $E=i\partial/\partial t$. Assume that the expression is an operator acting on $\psi(\vec{x},t)$. First

$$(\beta m+(E-eV))(\beta m-(E-eV))=m^2-(E-eV)^2\tag{2.145}$$

and then

$$\begin{aligned}(\vec{\alpha}\cdot(\vec{p}-e\vec{A}))&(\beta m-(E-eV))+(\beta m+(E-eV))(\vec{\alpha}\cdot(\vec{p}-e\vec{A}))\\&=ie\vec{\alpha}\cdot\vec{E},\end{aligned}\tag{2.146}$$

where $\vec{E}=-\vec{\nabla}V-\partial\vec{A}/\partial t$ is the electric field. Collecting terms, Equation (2.141) becomes:

$$[(\vec{p}-e\vec{A})^2-e\vec{\Sigma}\cdot(\vec{\nabla}\times\vec{A})+m^2-(E-eV)^2+ie\vec{\alpha}\cdot\vec{E}]\psi=0.\tag{2.147}$$

So far, no approximations. Now subtract the mass from the total energy: $E = E' + m$, where in the non-relativistic limit $E' \ll m$, and divide by $2m$ to obtain the expression:

$$\begin{aligned}&[(\vec{p} - e\vec{A})^2/(2m) - e\vec{\Sigma} \cdot \vec{B}/(2m) - E'^2/(2m) - E' \\ &\quad + 2(E' + m)(eV)/(2m) - (eV)^2/(2m) + ie\vec{\alpha} \cdot \vec{E}/(2m)]\psi = 0.\end{aligned} \tag{2.148}$$

In the non-relativistic limit drop $E'^2/(2m)$ and $(eV)^2/(2m)$ and replace $2(E' + m)/(2m) \approx 1$, but retain the explicitly relativistic term $ie\vec{\alpha} \cdot \vec{E}/(2m)$ (α mixes large and small components of ψ) to obtain

$$[(\vec{p} - e\vec{A})^2/(2m) - e\vec{\Sigma} \cdot \vec{B}/(2m) - E' + eV + ie\vec{\alpha} \cdot \vec{E}/(2m)]\psi = 0. \tag{2.149}$$

Operating on the column vector:

$$\psi = \begin{pmatrix} \psi_a \\ \psi_b \end{pmatrix} \tag{2.150}$$

gives

$$\begin{aligned}&[(\vec{p} - e\vec{A})^2/(2m) + eV - e\vec{\sigma} \cdot \vec{B}/(2m) - E']\psi_a \\ &\quad + (ie\vec{\sigma} \cdot \vec{E}/(2m))\psi_b = 0.\end{aligned} \tag{2.151}$$

The first four terms, operating on ψ_a, a two-component spin vector, are identical to the non-relativistic Schrödinger equation for a spin 1/2 electron in an electromagnetic field (remember that 'e' in these formulas is the positive elementary charge) described by $(V, \vec{A})$, with the important addition of a spin magnetic moment term, that has to be inserted by hand in the NR theory. Not only does the term $U = -\vec{\mu_e} \cdot \vec{B}$ appear automatically, but also the electron g factor is two. Recall that by definition for an elementary particle of charge q $\vec{\mu} = g \times q/(2m)\vec{S}$. Since $\vec{\sigma} = 2\vec{S}$, $g = 2$ falls out of the calculations as a prediction. $g = 2$ is a magic number, because it means that the spin and momentum precess in a perpendicular magnetic field at the same rate. The cyclotron frequency $\omega_c = eB/m$ is the same as the spin precession frequency $\omega_s = geB/2m$ if $g = 2$, so if the spin and momentum are parallel entering the magnetic field, they are still parallel upon exit. (Time dilation puts a factor of γ in each denominator.) A very sensitive technique for measuring $g-2$ is to observe the change in angle of the spin relative to the momentum after storage in a transverse magnetic field. The correct relativistic formula for the difference in precession frequencies is $\omega_s - \omega_c = \omega_c \gamma (g/2 - 1)$.

The last term in Equation (2.151) operating on ψ_b, is the spin orbit force. It is easy to show using $\psi_b = ((\vec{\sigma} \cdot \vec{p})\psi_a)/(E+m)$ that:

$$(ie\vec{\sigma} \cdot \vec{E}/(2m))\psi_b \approx [((Ze^2)/(2m^2 r^3))\vec{S} \cdot \vec{L}]\psi_a. \tag{2.152}$$

The derivation of this formula — the spin–orbit force for a single electron in orbit around a nucleus of charge $+Ze$ — is vexing in quantum mechanics class. One begins with the relativistic effect that a charged particle moving in the circular orbit of a radial electric field observes a transverse magnetic field: $\vec{B} = -\vec{v} \times \vec{E}$. Then couple the electron's magnetic moment to this field $U = -\vec{\mu_e} \cdot \vec{B}$, and do a little algebra to obtain $\vec{L} = \vec{r} \times \vec{p}$ from the expression for $\vec{B}$. One obtains a formula like Equation (2.152), except that it is a factor of two too large! The lecturer then mumbles something about the 'Thomas precession' that subtracts off half of the result, and gives the answer in agreement with experiment. The derivation of the Thomas precession, that arises because two non-collinear Lorentz transformations are equal to a third plus a space rotation, is rarely given in class, or in quantum mechanics texts. But Dirac's equation gets it right to start with, automatically including the effect of Mr. Thomas, a great benefit. Dirac's equation includes all relativistic effects.

There is an alternate derivation of the electron magnetic moment energy in an external magnetic field that does not use relativity. Sakurai attributes this to Feynman. The analysis is done with two component spin vectors, and the Pauli matrices. Since $(\vec{\sigma} \cdot \vec{p})(\vec{\sigma} \cdot \vec{p}) = \vec{p}^{\,2}$, one may write the kinetic energy thus,

$$\vec{p}^{\,2}/(2m) = (\vec{\sigma} \cdot \vec{p})(\vec{\sigma} \cdot \vec{p})/(2m), \tag{2.153}$$

where we have used the identity $(\vec{\sigma} \cdot \vec{a})(\vec{\sigma} \cdot \vec{b}) = \vec{a} \cdot \vec{b} + i\vec{\sigma} \cdot (\vec{a} \times \vec{b})$ that holds for $\vec{a}$ and $\vec{b}$ operators that commute with $\vec{\sigma}$ but not necessarily with each other. Nothing new so far, except that the wave function has two components. Now substitute $\vec{p} \rightarrow \vec{p} - e\vec{A}$, and remember that $\vec{p}$ and $\vec{A}$ do not commute. After the vector identity $\vec{\nabla} \times (\vec{A}\psi) = \psi(\vec{\nabla} \times \vec{A}) - \vec{A} \times (\vec{\nabla}\psi)$, the gauge condition $\vec{\nabla} \cdot \vec{A} = 0$, $\vec{B} = \vec{\nabla} \times \vec{A}$, and the choice $\vec{A} = (yB\hat{x} - xB\hat{y})/2$ one obtains:

$$\begin{aligned}&(\vec{\sigma} \cdot (\vec{p} - e\vec{A}))(\vec{\sigma} \cdot (\vec{p} - e\vec{A}))/(2m) \\ &\quad = \vec{p}^{\,2}/(2m) + (e\vec{A})^2/(2m) - e\vec{L} \cdot \vec{B}/(2m) - e\vec{\sigma} \cdot \vec{B}/(2m).\end{aligned} \tag{2.154}$$

The spin magnetic moment term is the same as in Equation (2.151), and no relativistic wave equation has been used. Some authors have argued that $g = 2$ is 'natural' for an elementary particle of arbitrary spin [7]. We shall revisit this question in a later chapter, when we consider the magnetic moments of the $W^\pm$ weak bosons. The orbital term has $g = 1$.

The hydrogen atom can be solved using the Dirac equation for the electron. One obtains two first-order coupled differential equations. The orbital angular momentum $\vec{L} = \vec{r} \times \vec{p}$ does not commute with a Hamiltonian which is linear in $\vec{p}$. Recall that the non-relativistic Hamiltonian involves $\vec{p}^{\,2}$, the angular part of which is $\vec{L}^2$. The energy levels depend on the principal quantum number n and the total angular momentum quantum number j, where $\vec{J} = \vec{L} + \vec{S}$. The 'exact' energy in terms of the electron mass, the fine structure constant α, the total angular momentum quantum number $j = l \pm 1/2$, and the principal quantum number n is:

$$E/m = (1 + \alpha^2/(((j + 1/2)^2 - \alpha^2)^{1/2} + n - j - 1/2)^2)^{-1/2}. \tag{2.155}$$

Expanding this tedious formula in a Taylor series in powers of α^2 gives to order α^4:

$$E = m + E_0 + 2E_0^2(3/4 - n/(j + 1/2))/m, \tag{2.156}$$

where $E_0 = -(m\alpha^2)/(2n^2) = -13.6\ \text{eV}/n^2$ is Bohr's formula for the energy levels of hydrogen. The fine structure constant α is a measure of the electron velocity in the hydrogen ground state, and $m\alpha^2/2$ is the kinetic energy. This same equation to order α^4 can be obtained by applying perturbation theory to Schrödinger's non-relativistic equation, evaluating two terms: the spin–orbit coupling, and a correction to the kinetic energy $-p^4/(8m^3)$. Equation (2.156) gives the fine structure splitting for the $2P^{3/2} - 2P^{1/2}$ levels of hydrogen as 4.5×10^{-5} eV, or about 1.1×10^{10} Hz, in agreement with experiment. These results — the electron g factor, the spin–orbit force, and the fine structure splitting in hydrogen — gave support that Dirac's linear equation was on the right track.

Higher powers of α^2 are irrelevant, since the exact formula predicts the same energy for the $2P^{1/2}$ and $2S^{1/2}$ levels. These levels are split by the Lamb shift of 1057 MHz, about 10% of the fine structure. The $2S^{1/2}$ level lies above the $2P^{1/2}$ level. The $2S^{1/2}$ is metastable, because the radiative decay $2S^{1/2} \to 1S^{1/2} + \gamma$ is strictly forbidden, and the direct decay $2S^{1/2} \to 2P^{1/2} + \gamma$ has a very slow rate because of the small energy difference. Metastables passing through a microwave cavity at 1057 MHz drop to the ground state via $2S^{1/2} + \gamma_1 \to 2P^{1/2} + 2\gamma_1$(stimulated emission) $\to 1S^{1/2} + \gamma_2$. The Lamb shift splitting is a radiative correction, beyond the scope of the single particle relativistic theory, but can be calculated using quantum electrodynamics, which treats the electron and photon as fields. Quantum field theory is therefore of fundamental importance in understanding nature.

In quantum mechanics, one learns that the first thing to do is diagonalize the Hamiltonian. $H = \vec{\alpha} \cdot \vec{p} + m\beta$ can be diagonalized, with eigenvalues $+|E|$ and $-|E|$ each occurring twice, with ψ_a belonging to $+|E|$, and ψ_b belonging to $-|E|$. This form is never used, however. It is interesting that the equation in its original form has much more physical content than does its diagonal version. The presence of two eigenvalues $\pm|E|$ emphasizes the fact that the negative energies cannot be ignored, and have to be interpreted in a satisfactory manner. Dirac realized the problem early, and proposed that all of the negative energy states were filled by electrons, so that the Pauli exclusion principle would forbid any transition from positive to negative energy. The absence of a negative energy electron — a hole — would act as a positively charged particle, hence predicting the existence of positrons. The vacuum would carry infinite charge. There are some parallels with semiconductors, where the absence of an electron acts like a positive carrier, but the semiconductor itself is electrically neutral.

There are other anomalies of the single particle Dirac equation that are fully discussed in Sakurai's book [5] cited earlier. We have derived $\langle \alpha_z \rangle$ for positive and negative energy spinors separately, and obtained reasonable results. Sakurai shows that, while for free particle plane waves the wave function may contain only positive energies, if the wave packet is localized to dimensions less than the Compton wave length $1/m$, then positive and negative energies are both required to explain the development of the packet with time. The non-relativistic formula for the change in time of a minimum uncertainty Gaussian wave packet is presented in Equation (3.67), where difficulties with the speed of light are clearly displayed. Calculation of $\langle \alpha_z \rangle$, or any other observable, for a wave function with both positive and negative energies results in an expression that has a rapid oscillatory time dependence. This curiosity was given the German name *Zitterbewegung* by Schrödinger, translated by Sakurai as 'quivering motion'.

Then there is the Klein paradox. Take a positive energy electron moving in a constant positive potential V. The Dirac equation for the motion of the upper two components ψ_a can be written:

$$\left(\vec{\sigma} \cdot \vec{p} \frac{1}{((E+m)-V)} \vec{\sigma} \cdot \vec{p}\right) \psi_a = ((E-m)-V)\psi_a, \tag{2.157}$$

which, since the denominator is a constant, gives the relation

$$p^2 = ((E+m)-V)((E-m)-V). \tag{2.158}$$

Positive p^2 gives oscillations, and negative p^2 gives exponential damping-barrier penetration. All of the numbers, $E, m,$ and V are positive. $V = 0$ gives $p^2 = E^2 - m^2 > 0$. If $V > E - m$ and $V < E + m$, then $p^2 < 0$ — exponential damping. However, if $V > E + m$, then both terms $((E+m)-V)$ and $((E-m)-V)$ are negative, and $p^2 > 0$, so it oscillates again! This is Klein's paradox.

The field theory treatment of the Dirac equation, with electrons moving forward in time, and positrons moving backward, both with positive energy, avoids the question of interpretation of the filled negative energy sea.

Positron spinors are defined as negative energy $E \to -|E|$, negative space momentum $\vec{p} \to -\vec{p}$, and flipped spin:

$$u^{(3,4)}(-p)e^{-i(-p)\cdot x} = v^{(2,1)}(p)e^{ip\cdot x}, \tag{2.159}$$

where $u^{(3,4)}$ are given in Equation (2.129). Explicitly the positron spinors are

$$v^{(1)}(p) = \sqrt{E+m}\begin{pmatrix} (p_x - ip_y)/(E+m) \\ -p_z/(E+m) \\ 0 \\ 1 \end{pmatrix};$$

$$v^{(2)}(p) = \sqrt{E+m}\begin{pmatrix} p_z/(E+m) \\ (p_x + ip_y)/(E+m) \\ 1 \\ 0 \end{pmatrix}. \tag{2.160}$$

The positron energy is to be treated as a positive quantity. Normalizations are

$$v^{(1)\dagger}(p) \times v^{(1)}(p) = v^{(2)\dagger}(p) \times v^{(2)}(p) = 2E, \tag{2.161}$$

and

$$\bar{v}^{(1)}(p) \times v^{(1)}(p) = \bar{v}^{(2)}(p) \times v^{(2)}(p) = -2m. \tag{2.162}$$

Cross terms are orthogonal as before. There is more than one subtlety in treatment of positron spinor algebraic signs. For instance, while $v^{(s)}(p)^{\dagger}$ is not orthogonal to $u^{(s')}(p)$, if you reverse the space momentum vector to $v^{(s)}(-\vec{p})$, they *are* orthogonal. Because the four momentum has been reversed, the Dirac equation for a positron spinor reads:

$$(\not{p} + m)v(p) = 0. \tag{2.163}$$

The adjoint equation is

$$\bar{v}(p)(\not{p} + m) = 0. \tag{2.164}$$

Completeness sums for electron and positron spinors are

$$\sum_{s=1,2} u^{(s)}(p)\bar{u}^{(s)}(p) = (\not{p} + m), \tag{2.165}$$

$$\sum_{s=1,2} v^{(s)}(p)\bar{v}^{(s)}(p) = (\not{p} - m). \tag{2.166}$$

These expressions will be useful calculational aids later.

2.8. Lorentz Transformation of Dirac's Equation

The Dirac equation in covariant form is

$$\gamma^{\mu}\partial\psi(x)/\partial x^{\mu} + im\psi(x) = 0. \tag{2.167}$$

Under Lorentz transformation, the coordinates transform thus,

$$x'^{\mu} = L^{\mu}_{\nu}x^{\nu}, \tag{2.168}$$

where the matrix elements are $L^0_0 = L^3_3 = \gamma, L^0_3 = L^3_0 = v\gamma$, $L^1_1 = L^2_2 = 1$, and all other elements are zero. We write the Dirac equation in the x' frame as

$$\gamma^{\mu}\partial\psi'(x')/\partial x'^{\mu} + im\psi'(x') = 0. \tag{2.169}$$

Note that the γ^{μ} matrices are the same in the two frames. We seek a matrix $S(v, \gamma)$ that is not a function of the coordinates, and performs the following operation:

$$\psi'(x') = S\psi(x). \tag{2.170}$$

The matrix S is 4×4, and does not necessarily commute with the matrices γ^{μ}. The inverse $S^{-1}\psi'(x') = \psi(x)$ must also exist. Multiply Equation (2.169) from the left by S^{-1}, insert SS^{-1} between γ^{μ} and $\psi'(x')$, and Lorentz transform $\partial/\partial x'^{\mu}$ to obtain

$$S^{-1}\gamma^{\mu}SL^{\nu}_{\mu}\partial\psi(x)/\partial x^{\nu} + im\psi(x) = 0. \tag{2.171}$$

Compare with Equation (2.167) to obtain the formula:

$$S\gamma^{\nu}S^{-1} = L^{\nu}_{\mu}\gamma^{\mu}. \tag{2.172}$$

Remember that S does not commute with $\psi(x)$, and L^{ν}_{μ} is a number. Substituting in the matrix elements for our Lorentz transformation gives:

$$S\gamma^0 S^{-1} = \gamma\gamma^0 + v\gamma\gamma^3, \tag{2.173}$$

$$S\gamma^1 S^{-1} = \gamma^1, \tag{2.174}$$

$$S\gamma^2 S^{-1} = \gamma^2, \tag{2.175}$$

$$S\gamma^3 S^{-1} = v\gamma\gamma^0 + \gamma\gamma^3. \tag{2.176}$$

Hopefully the two uses of the Greek letter γ does not cause confusion. The two middle equations show that S commutes with both γ^1 and γ^2, which suggests a form $S = a + b\gamma^0\gamma^3$. A little algebra gives:

$$a/b = (v\gamma)/(\gamma - 1), \tag{2.177}$$

$$S = \sqrt{(\gamma+1)/2} + \sqrt{(\gamma-1)/2}\gamma^0\gamma^3, \tag{2.178}$$

$$S^{-1} = \sqrt{(\gamma+1)/2} - \sqrt{(\gamma-1)/2}\gamma^0\gamma^3. \tag{2.179}$$

The explicit form for the matrix S is

$$S = \sqrt{(\gamma+1)/2} \times \begin{pmatrix} 1 & 0 & \sqrt{\frac{\gamma-1}{\gamma+1}} & 0 \\ 0 & 1 & 0 & -\sqrt{\frac{\gamma-1}{\gamma+1}} \\ \sqrt{\frac{\gamma-1}{\gamma+1}} & 0 & 1 & 0 \\ 0 & -\sqrt{\frac{\gamma-1}{\gamma+1}} & 0 & 1 \end{pmatrix}. \tag{2.180}$$

This is the matrix S for a proper Lorentz transformation. For time reversal, $L^0_0 = -1$; $L^1_1 = L^2_2 = L^3_3 = +1$. All other elements vanish. For time reversal, $S_T = \gamma^1\gamma^2\gamma^3 = S_T^{-1}$. While proper S involves the product of two γ matrices, for time reversal S_T involves three, and for space inversion (parity — that reverses the space coordinates but not the time), $S_P = S_P^{-1}$ is just γ^0. For time reversal in our representation for the γ matrices, the matrix is

$$\text{For time reversal, } S_T = i\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{2.181}$$

This matrix changes the spinor $u^{(1)}(p) \to iu^{(3)}(p)$. An important consequence of the parity operation $S_P = \gamma^0$ is that the upper and lower components ψ_a and ψ_b of the spinor have opposite parities. Since the positrons are associated with the lower two components, this means that positronium in its ground state has odd intrinsic parity, and if the quarks are elementary Dirac spin 1/2 particles, so do the ground states of $c\bar{c}$ and $b\bar{b}$.

Equation (2.172) is the key to Lorenz transformations of expectation values of the matrices γ^μ. Other useful properties of S are

$$S^\dagger = S, \tag{2.182}$$

$$S^{-1} = \gamma^0 S \gamma^0, \tag{2.183}$$

$$\psi'(x') = S\psi(x), \tag{2.184}$$

$$\bar{\psi}'(x') = \bar{\psi}(x)S^{-1}. \tag{2.185}$$

The last two equations show that $\bar{\psi}(x)\psi(x) = \bar{\psi}'(x')\psi'(x')$, a Lorentz invariant scalar. Combined with Equation (2.172), we have:

$$\bar{\psi}'(x')\gamma^\mu\psi'(x') = L^\mu_\nu\bar{\psi}(x)\gamma^\nu\psi(x). \tag{2.186}$$

So that $\bar{\psi}(x)\gamma^\mu\psi(x)$ Lorentz transforms like the coordinates — a four-vector. A tensor can be formed from the product $\gamma^\mu\gamma^\nu$:

$$\gamma^\mu\gamma^\nu = (\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)/2 + (\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)/2 \tag{2.187}$$

or

$$\gamma^\mu\gamma^\nu = g^{\mu\nu} + (\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)/2. \tag{2.188}$$

The metric tensor is Lorentz invariant, so we are left with the totally antisymmetric part that has six independent components: $\gamma^0\gamma^1, \gamma^0\gamma^2, \gamma^0\gamma^3, \gamma^1\gamma^2, \gamma^1\gamma^3, \gamma^2\gamma^3$. We define

$$\sigma^{\mu\nu} = i(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)/2, \tag{2.189}$$

$$\bar{\psi}'(x')\sigma^{\mu\nu}\psi'(x') = L^\mu_\alpha L^\nu_\beta \bar{\psi}(x)\sigma^{\alpha\beta}\psi(x). \tag{2.190}$$

There are four 4×4 matrices γ^μ with $\mu = 0, 1, 2, 3$, defined as $\gamma^0 = \beta$, $\gamma^j = \beta\alpha^j$. There is a fifth matrix that anticommutes with all of the γ^μ, called appropriately enough γ^5:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = (i/4!)\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma, \tag{2.191}$$

$$\gamma^{5\dagger} = \gamma^5, \tag{2.192}$$

$$(\gamma^5)^2 = 1, \tag{2.193}$$

$$\gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0,\ \mu = 0, 1, 2, 3, \tag{2.194}$$

$$\gamma^5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{2.195}$$

An important spin projection operator for massless spin 1/2 particles like neutrinos is $(1 \pm \gamma^5)/2$ which are the matrices:

$$(1 \pm \gamma^5)/2 = 1/2 \begin{pmatrix} 1 & 0 & \pm 1 & 0 \\ 0 & 1 & 0 & \pm 1 \\ \pm 1 & 0 & 1 & 0 \\ 0 & \pm 1 & 0 & 1 \end{pmatrix}. \tag{2.196}$$

A massless Dirac particle has no rest frame. With momentum along the $\hat{z}$ direction and $\langle\sigma_z\rangle = \pm 1$, the spinors are

$$u^{(1)}(p) = \sqrt{E}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}; \quad u^{(2)}(p) = \sqrt{E}\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}. \tag{2.197}$$

Obviously $(1+\gamma^5)/2u^{(1)}(p) = u^{(1)}(p)$, and $(1-\gamma^5)/2u^{(2)}(p) = u^{(2)}(p)$, while $(1-\gamma^5)u^{(1)}(p) = (1+\gamma^5)u^{(2)}(p) = 0$. These projection operators satisfy the rules:

$$\begin{aligned} &((1+\gamma^5)/2)^2 = (1+\gamma^5)/2,\ ((1-\gamma^5)/2)^2 = (1-\gamma^5)/2, \\ &(1+\gamma^5)/2 + (1-\gamma^5)/2 = 1,\ (1+\gamma^5)(1-\gamma^5) = 0. \end{aligned} \tag{2.198}$$

For massless positrons, or anti-neutrinos, the roles are reversed. The spinors are

$$v^{(1)}(p) = \sqrt{E}\begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \quad v^{(2)}(p) = \sqrt{E}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

$$\frac{(1-\gamma^5)}{2}v^{(1)} = v^{(1)}, \quad \frac{(1+\gamma^5)}{2}v^{(2)} = v^{(2)}. \tag{2.199}$$

The index for γ^5 cannot be lowered by the metric tensor — it has no ‘5’ index! For the record, $i\gamma_0\gamma_1\gamma_2\gamma_3 = -\gamma^5$. In our representation for the matrices $\vec{\alpha}$ and β, γ^5 mixes large and small components of $\psi(x)$. $\bar{\psi}(x)\gamma^5\psi(x)$ has some unusual properties. Proper Lorentz transformations are bilinear in the γ^μ’s, and commute with γ^5, but improper Lorentz transformations — parity space inversion and time reversal — involve an odd number of γ^μ’s, and therefore anticommute with γ^5. To summarize

$$\bar{\psi}'(x')\gamma^5\psi'(x') = +\bar{\psi}(x)\gamma^5\psi(x), \tag{2.200}$$

for a proper Lorentz transformation; and

$$\bar{\psi}'(x')\gamma^5\psi'(x') = -\bar{\psi}(x)\gamma^5\psi(x), \tag{2.201}$$

for space inversion or time reversal. The expression $\bar{\psi}(x)\gamma^5\psi(x)$ is called a pseudoscalar. In three dimensions, the orbital angular momentum $\vec{L} = \vec{r} \times \vec{p}$ is an axial vector, as is the magnetic field $\vec{\nabla} \times \vec{B} = \mu_0\vec{J}$. Vectors, like $\vec{r}$, or $\vec{J}$ change sign under space inversion, while axial vectors that are defined by a circulation loop, do not. The extension of a cross product to four dimensions gives a six-vector, like $\sigma^{\mu\nu}$, but it is the product $\gamma^5\gamma^\mu$ that forms what is called an axial vector in relativistic quantum mechanics. The invariants are summarized in the table.

Name	Form	Number
Scalar	$\bar{\psi}\psi$	1
Pseudoscalar	$\bar{\psi}\gamma^5\psi$	1
Vector	$\bar{\psi}\gamma^\mu\psi$	4
Axialvector	$\bar{\psi}\gamma^5\gamma^\mu\psi$	4
Tensor	$\bar{\psi}\sigma^{\mu\nu}\psi$	6
Total		16

Note that the 16 expressions cover the space of 4×4 'Dirac Hermitian' matrices $M^\dagger = \gamma^0 M \gamma^0$. For this γ^5 needs an i.

2.9. Gordon Decomposition of the Vector Current

The vector current $j^\mu = e\bar{\psi}\gamma^\mu\psi$ plays a special role in quantum electrodynamics, because the current density is the source of the vector potential. The decomposition into two terms, named after Walter Gordon, separates the current into a part that is only the charge density and a part that is the magnetic moment density. The first part has no γ matrices. Of special interest is the 'transition current density' between two states $u(p_i)$ and $u(p_f)$:

$$e\bar{u}(p_f)\gamma^\mu u(p_i) = e\left\{\bar{u}(p_f)((p_f + p_i)^\mu + i\sigma^{\mu\nu}(p_f - p_i)_\nu)u(p_i)\right\}/(2m), \tag{2.202}$$

where

$$\sigma^{\mu\nu} = i(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)/2. \tag{2.203}$$

The simplest way to demonstrate this expression is to work from the right hand side, since we already know what it is, and use the Dirac equation. To wit,

$$\begin{aligned} &-\bar{u}_f(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)(p_f - p_i)_\nu u_i/2 \\ &\quad = -(\bar{u}_f\gamma^\mu \not{p}_f u_i - \bar{u}_f\gamma^\mu \not{p}_i u_i - \bar{u}_f \not{p}_f \gamma^\mu u_i + \bar{u}_f \not{p}_i \gamma^\mu u_i)/2. \end{aligned} \tag{2.204}$$

The second and third terms each give $-m\bar{u}_f\gamma^\mu u_i$. For the first and fourth terms the γ^μ is on the wrong side. This is easy to fix with the aid of the anticommutation rule $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$. Thus,

$$\bar{u}_f\gamma^\mu \not{p}_f u_i = \bar{u}_f\gamma^\mu\gamma^\nu p_{f\nu}u_i = -\bar{u}_f\not{p}_f\gamma^\mu u_i + 2\bar{u}_f u_i p_f^\mu. \tag{2.205}$$

Assembly of all the parts proves the theorem. Note that the charge current density involves the sum of the two four momenta, while the current density due to the electron magnetic moment is multiplied by the difference — the momentum transfer.

2.10. Charge Conjugation and TCP

Charge conjugation is an operation that changes electrons into positrons. Parity and time reversal have been treated here as Lorentz transformations, but charge conjugation changes the sign of the four momentum $p^\mu \to -p^\mu$ without also changing the sign of x^μ, and that is not an operation that can be accomplished by a Lorentz transformation. While it is possible to change the spinor $u(p) \to v(p)$ with time reversal (see Problem 2.26), it is not possible to change the phase $e^{-ip^\mu x_\mu} \to e^{ip^\mu x_\mu}$, because $p^\mu x_\mu$ is Lorentz invariant. However, the phase of the wave function can be reversed by complex conjugation, so we may define the charge conjugate wave function thus,

$$v^{(1)}(p) = C \times u^{(1)}(p)^* = i\gamma^2 u^{(1)}(p)^*, \tag{2.206}$$

or

$$\begin{pmatrix} (p_x - ip_y)/(E+m) \\ -p_z/(E+m) \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ p_z/(E+m) \\ (p_x - ip_y)/(E+m) \end{pmatrix}. \tag{2.207}$$

The same matrix changes $u^{(2)}(p)^*$ into $-v^{(2)}(p)$. Thus,

$$v^{(2)}(p) = -i\gamma^2 u^{(2)}(p)^*. \tag{2.208}$$

The product of our three symmetry operations PTC on a Dirac spinor is

$$\mathrm{PTC} = i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^2 = \gamma^5\gamma^2. \tag{2.209}$$

This operator and complex conjugation flips the spin of an electron spinor. See Problem 2.28. The reason for writing out the product is that quantum field theories obey a general TCP theorem, that states that any physical system is invariant under the product of these three operations in any order. We will often use the general theorem without formal proof. That electrodynamics is invariant under each operation separately is consistent with experiment, but this is not true for the weak interactions. CP violation is discussed in Chapter 9, where the TCP theorem is used to argue that CP violation necessarily also means T violation, or in other words CP *is* T. In this regard it is worth noting that the Lorentz transformation operator S_T (Equation (2.180)) changes the spinor $u^{(1)}(p) \to iu^{(3)}(p) = iv^{(2)}(-p)$, where $v^{(2)}(p)$ is a positron spinor, defined in Equation (2.159).

2.11. Problems

2.1 Show that in the expression for a rotation through three Euler angles about sequential axes:

$$R(\alpha,\beta,\gamma) = e^{-i\gamma L_{z''}} \times e^{-i\beta L_{y'}} \times e^{-i\alpha L_z} \tag{2.210}$$

can be written in terms of fixed axes if the order is reversed:

$$R(\alpha,\beta,\gamma) = e^{-i\alpha L_z} \times e^{-i\beta L_y} \times e^{-i\gamma L_z}. \tag{2.211}$$

Thus, you can erase the primes if you invert the order. The simplest way to do this is use $e^{-i\gamma L_{z''}} = e^{-i\beta L_{y'}} \times e^{-i\gamma L_z} \times e^{i\beta L_{y'}}$, and $L_{z'} = L_z$.

2.2 The simplest rotation matrices are for spin 1/2. Because $\vec{S} = \vec{\sigma}/2$, the rotation matrices involve 1/2 angles. Show that:

$$e^{-i\alpha S_z} = \cos(\alpha/2) - i\sigma_z \sin(\alpha/2). \tag{2.212}$$

This formula holds for all three components. Obtain the rotation matrix for Euler angles α and β:

$$\begin{aligned} R(\alpha,\beta) &= e^{-i\alpha\sigma_z/2} \times e^{-i\beta\sigma_y/2} \\ &= \begin{pmatrix} e^{-i\alpha/2}\cos(\beta/2) & -e^{-i\alpha/2}\sin(\beta/2) \\ e^{i\alpha/2}\sin(\beta/2) & e^{i\alpha/2}\cos(\beta/2) \end{pmatrix}. \end{aligned} \tag{2.213}$$

Since $e^{i\pi} = -1$, this operation gives a minus sign for rotation about the z axis through 360°. The two component spin function χ transforms as $R(2\pi)\chi = -\chi$.

2.3 The next rotation matrices are for $l = 1$. Set the Euler angle $\gamma = 0$. For the Cartesian coordinates (x, y, z) the 3×3 rotation matrix is

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} \cos\beta\cos\alpha & \cos\beta\sin\alpha & -\sin\beta \\ -\sin\alpha & \cos\alpha & 0 \\ \sin\beta\cos\alpha & \sin\beta\sin\alpha & \cos\beta \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{2.214}$$

The usual polar coordinates are

$$x = r\sin\theta\cos\phi, \tag{2.215}$$

$$y = r\sin\theta\sin\phi, \tag{2.216}$$

$$z = r\cos\theta. \tag{2.217}$$

The spherical harmonics for $l = 1$ are

$$Y_1^{+1} = -(3/(8\pi))^{1/2}\sin\theta e^{i\phi}, \tag{2.218}$$

$$Y_1^0 = (3/(4\pi))^{1/2}\cos\theta, \tag{2.219}$$

$$Y_1^{-1} = (3/(8\pi))^{1/2}\sin\theta e^{-i\phi}. \tag{2.220}$$

Obtain the mixing matrix that transforms the three spherical harmonics into Cartesian coordinates:

$$x = (4\pi/3)^{1/2} rMY. \tag{2.221}$$

Show that M is unitary: $M^{-1} = M^{\dagger}$; and derive the $l = 1$ rotation matrix elements from $M^{\dagger}RM$, where R is the Cartesian matrix above.

2.4 Another way to obtain the rotation matrix elements is to use the operator definition $R(\alpha, \beta) = e^{-i\alpha L_z} \times e^{-i\beta L_y}$, and the form for $l = 1$ used in Problem 1.4 on G-parity:

$$e^{-i\beta L_y} = 1 + (\cos\beta - 1)L_y^2 - i\sin\beta L_y, \tag{2.222}$$

and

$$e^{-i\alpha L_z} = 1 + (\cos\alpha - 1)L_z^2 - i\sin\alpha L_z. \tag{2.223}$$

The second equation is easy to handle because both L_z and L_z^2 are diagonal. Multiply the matrices together to give the matrix: $D^1_{m',m}(\alpha, \beta)$, which is the transpose of the rotation matrix of Problem 2.3, namely,

$$D^1_{m',m}(\alpha,\beta) = \begin{pmatrix} e^{-i\alpha}(1+\cos\beta)/2 & -e^{-i\alpha}\sin\beta/\sqrt{2} & e^{-i\alpha}(1-\cos\beta)/2 \\ \sin\beta/\sqrt{2} & \cos\beta & -\sin\beta/\sqrt{2} \\ e^{i\alpha}(1-\cos\beta)/2 & e^{i\alpha}\sin\beta/\sqrt{2} & e^{i\alpha}(1+\cos\beta)/2 \end{pmatrix}. \tag{2.224}$$

The matrices are transposed because of the convention for the D matrices:

$$Y_1^m(\theta', \phi') = \sum_{m'=-1}^{+1} Y_1^{m'}(\theta, \phi) D^1_{m',m}(\alpha, \beta, \gamma). \tag{2.225}$$

2.5 Calculate the determinant of the 'proper' Lorentz transformation Equation (2.25). Write down the Lorentz transformation for time reversal $t \to -t$ and no other changes. Then write down the Lorentz transformation for space inversion $\vec{x} \to -\vec{x}$ and no other changes. Calculate these determinants. They are called 'improper' Lorentz transformations.

2.6 **Rapidity.** The Lorentz transformation parameters $\vec{v}$ and $\gamma = 1/\sqrt{1 - v^2}$ are related to hyperbolic functions. Show that in terms of the angle χ, called the rapidity, one can use the trigonometric identities for hyperbolic functions to write:

$$v = \tanh \chi, \tag{2.226}$$

$$\gamma = \cosh \chi, \tag{2.227}$$

$$v\gamma = \sinh \chi. \tag{2.228}$$

Show that the hyperbolic functions have the correct ranges: $1 \leq \gamma \leq \infty$ and $-1 \leq v \leq +1$. While in relativity velocities do not simply add, rapidities do. For two collinear Lorentz transformations, $\chi = \chi_1 + \chi_2$. Show that Einstein's velocity addition theorem $v = (v_1 + v_2)/(1 + v_1 v_2)$ follows from the rules for combining hyperbolic functions. Express the Lorentz transformation in terms of χ.

2.7 **More rapidity.** Lorentz transformations by convention are normally along the z axis. Momentum components perpendicular to the z axis, called $p_\perp$, are Lorentz invariant, while parallel components, $p_{\|}$ change from one moving frame to the next. It is convenient to introduce a variation of the rapidity χ, called y, defined as

$$y = (1/2) \ln\big((E + p_{\|})/(E - p_{\|})\big). \tag{2.229}$$

While it may seem confusing, this definition is also called rapidity. Show that

$$\tanh y = p_{\|}/E, \tag{2.230}$$

$$\cosh y = E/\sqrt{E^2 - p_{\|}^2}. \tag{2.231}$$

Since $E^2 = m^2 + p_\perp^2 + p_{\|}^2$, the denominator in $\cosh y$ is often called the 'transverse mass'. Like the mass of a particle, the transverse mass is Lorentz invariant. Rapidities add, and differences in rapidity — rapidity gaps — are Lorentz invariant. If the total energy is $\sqrt{s}$, the maximum energy of a pion is $\sqrt{s}/2$, because there has to be another pion to balance the momentum. The maximum value of $p_{\|} = \pm\sqrt{s/4 - m^2}$. Show that for $\sqrt{s} \gg m$, $y_{max} = \pm \ln(\sqrt{s}/m)$. Calculate the full spread of rapidity for pions with $\sqrt{s} = 2\,\text{TeV}$.

2.8 Still more rapidity. Pseudorapidity $\eta = (1/2)\ln((1+\cos\theta)/(1-\cos\theta)) = -\ln(\tan(\theta/2))$ is y for a massless particle. η diverges as $\theta \to 0$. Detectors are often segmented in $|\Delta\eta|$, because this results in constant particle flux per segment. How many towers would cover $\theta = 90^\circ$ to $\theta = 20^\circ$ in steps $\Delta\eta = 0.1$? The towers are also segmented in $\Delta\phi$ in the transverse plain.

2.9 Two-body decay kinematics. A two-body decay like $K_s^0 \to \pi^+ + \pi^-$, $K^+ \to \mu^+ + \nu_\mu$, or $\pi^0 \to \gamma + \gamma$ is completely determined in the parent rest frame by the masses $m \to m_1 + m_2$. The momenta are equal and opposite, and the energies are given by Equation (2.63): $E_1 = (m^2 + m_1^2 - m_2^2)/(2m)$. The same formula with m_1 and m_2 interchanged gives E_2. If $m_1 = m_2$ then $E_1 = E_2 = m/2$. The momenta can be calculated from the energies, or from Equation (2.64). If the parent has no spin (like π's and K's), the distribution of momentum vectors $\vec{p_1}$ is isotropic in the parent rest frame, which is flat in $\cos\theta_c$. Show that an isotropic angular distribution in the parent rest frame maps into a uniform distribution in energy in the laboratory. Calculate the maximum and minimum energies for particle #1 in the laboratory — the ends of the box distribution.

2.10 If the parent is moving in the laboratory, $\vec{p_1}$ and $\vec{p_2} = -\vec{p_1}$ must be Lorentz transformed to give the final state momenta. Equation (2.62) gives the polar angle transformation formula for particle #1. The azimuthal angle is Lorentz invariant. For particle #2 $\cos\theta_c \to -\cos\theta_c$, and if $m_2 \neq m_1$ the parameter $a = v_c E_2/p_2$ has to be adjusted accordingly. One parameter of experimental interest is the opening angle between the two final state particles in the lab frame. Use the trigonometric formula for adding tangents to obtain the expression for $m_1 = m_2$:

$$\tan(\theta_1 + \theta_2) = (2a\gamma\sin\theta_c)/(\gamma^2(a^2 - \cos^2\theta_c) - \sin^2\theta_c). \tag{2.232}$$

Show that this formula has two possible extremes $d\tan(\theta_1 + \theta_2)/d\theta_c = 0$:

$$\cos\theta_c = 0, \quad \text{or} \quad \sin^2\theta_c = (\gamma^2(\mathrm{a}^2 - 1))/(\gamma^2 - 1). \tag{2.233}$$

The second condition can be satisfied if $a > 1$. This can be true for $K \to \pi\pi$, but not for $\pi^0 \to \gamma + \gamma$.

2.11 $\pi^0 \to \gamma + \gamma$. Here $a < 1$ because the π^0 velocity is always less than the γ. Calculate the energy of the γ emitted backwards in the laboratory for a 1.35 GeV π^0 ($\gamma = 10$). Plot the opening angle distribution vs. $\cos\theta_c$ for γ's from 1.35 GeV π^0's. Although not strictly correct, you may set $a = 1$ in the angle formula. Rather than a maximum value — ($\theta_1 + \theta_2 = \pi$ is always possible), the opening angle distribution has a minimum value $\theta_1 + \theta_2 \approx 2/\gamma$. The opening angle distribution of pairs of γ's is sharply peaked near the minimum — the γ's lie on a cone. A simple way to obtain the minimum opening angle is to look at a pair of γ's at $\theta_c = \pi/2$. Invariant transverse momentum $p_\perp = m/2$ and in the lab frame $p_\| = \gamma m/2$, so the opening angle is $2/\gamma$.

2.12 Calculate v_c for either pion in the decay $K^0_s \to \pi^+ + \pi^-$. The parameter a can be any number $0 \leq a \leq 1/v_c$. For K_0 beam momentum of 500 MeV/c plot a curve of $\theta_1 + \theta_2$ vs. $\cos\theta_c$. The opening angle distribution should have a minimum value (about 79°, coming from $\theta_c = \pi/2$). Repeat the calcultaion of the opening angle distribution for 1500 MeV/c K_0's. In this case the distribution has a maximum opening angle of about 33° that does not occur for $\theta_c = \pi/2$, but rather for an asymmetric two pion decay in the K rest frame. Even simple two-body kinematics has some surprises!

2.13 **Three-body decays.** An easy one to start with is $K^+ \to \pi^+\pi^+\pi^-$; all three masses are equal. Calculate the total kinetic energy available, the maximum energy of any pion, and the energy of pion #1 or #2 if #3 is at rest. The boundary curve lies within a square on a (x, y) plot, where $x = (E_1 - m_1)/(E_1^{\max} - m_1)$, and $y = (E_2 - m_2)/(E_2^{\max} - m_2)$; $0 \leq x \leq 1$, and $0 \leq y \leq 1$. The boundary curve is an oval inscribed in the square. Set $|\cos\theta_{12}| = 1$ in Equation (2.67), square both sides, and solve for the boundary curve $y = f(x)$.

2.14 An alternate form of the kinematic plot is sometimes convenient if the three final state particles all have the same mass.

The plot uses the fact that the sum of the three perpendiculars from any point inside an equilateral triangle is a constant. That constant is the total kinetic energy available, and the three perpendiculars are the kinetic energies of the particles. Show that the non-relativistic kinematic boundary is the inscribed circle. Obtain the boundary for $K \rightarrow \pi\pi\pi$. Show that the boundary curve for the rare decay $\rho^0 \rightarrow \gamma\gamma\gamma$ is an inverted inscribed equilateral triangle.

2.15 $K^+ \rightarrow \mu^+\pi^0\nu_\mu$. Here one particle, the neutrino, has zero mass, and the muon and pion have maximum energy simultaneously. Calculate $E_\mu^{\max}$ and $E_\pi^{\max}$. Solve Equation (2.69) for the boundary curve, and plot it on an (x, y) plot, where $x = (E_\mu - m_\mu)/(E_\mu^{\max} - m_\mu)$, and $y = (E_\pi - m_\pi)/(E_\pi^{\max} - m_\pi)$. This reproduces Figure 2.2.

2.16 Repeat 2.15 for $K^+ \rightarrow e^+\pi^0\nu_e$. Here both the positron and the neutrino can be treated massless. This reproduces Figure 2.3.

2.17 Relativistic wave equations. Show that the charge density ρ defined by Equation (2.91) reduces to $\rho = \psi^*\psi$ in the limit $E - m \ll m$, with the relativistic wave function $\phi(x, t) = \psi(x, t)e^{-imt}$. The non-relativistic wave function has the mass subtracted off from the energy.

2.18 Show that the 4×4 Dirac matrices written in 2×2 boxes have all the proper characteristics:

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.234}$$

Namely, $\alpha_i\alpha_j + \alpha_j\alpha_i = 2\delta_{i,j}$, $\beta\alpha_j + \alpha_j\beta = 0$, $\beta^2 = 1$, $(\vec{a} \cdot \vec{\alpha})(\vec{b} \cdot \vec{\alpha}) = \vec{a} \cdot \vec{b} + i\Sigma \cdot (\vec{a} \times \vec{b})$. The matrix $\vec{\Sigma}$ is defined in Equation (2.143).

2.19 Show that the spinors $u^{(s)}(p)$ with $s = 1, 2, 3, 4$ are consistent with Dirac's equation; $s = 1, 2$ for $E > 0$ and $s = 3, 4$ for $E < 0$. Check the normalizations and orthogonality for both $u^\dagger(p)u(p)$ and $\bar{u}(p)u(p)$.

2.20 Diagonalize the Hamiltonian $H = \vec{\alpha} \cdot \vec{p} + \beta m$.

2.21 Check the algebra in reducing Equations (2.141)–(2.147).

2.22 Show that the last term in Equation (2.149), $(-ie\vec{\alpha} \cdot \vec{E})\psi/(2m)$ reduces to the spin–orbit force of Equation (2.148).

2.23 Verify the completeness relations Equations (2.165) and (2.166) by multiplying from the right by $u^{(r)}(p)$ or $v^{(r)}(p)$, using the normalization, and the Dirac equation.

2.24 Show that the matrix $S = \cosh(\chi/2) + \sinh(\chi/2)\gamma^0\gamma^3$, where the rapidity χ is defined in Problem 2.6: $\tanh\chi = v$. You will need some half angle identities for hyperbolic functions: $\sinh x = 2\sinh(x/2)\cosh(x/2)$, just like trigonometry, and $\cosh x = \cosh^2(x/2) + \sinh^2(x/2)$, while $\cosh^2 x - \sinh^2 x = 1$. It was noted above that three-dimensional rotations for spin 1/2 vectors involve half angles. It is interesting that the Lorentz transformations for Dirac spinors also involve half angles — only the angle is rapidity, and the functions are hyperbolic.

2.25 Use the matrix for S given by Equation (2.180) to 'Lorentz boost' the rest frame spinor:

$$u^{(1)}(m) = \sqrt{2m}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}. \tag{2.235}$$

The final spinor has momentum $p_z = v\gamma m$ Even the normalization comes out correctly!

2.26 Show that the time reversal operation Equation (2.181) operating on the positive energy spinors give the positron spinors, namely,

$$S_T u^{(1)}(\vec{p}) = iv^{(2)}(-\vec{p}), \tag{2.236}$$

and

$$S_T u^{(2)}(\vec{p}) = iv^{(1)}(-\vec{p}). \tag{2.237}$$

2.27 Prove the Gordon decomposition of the vector current, Equation (2.202), using the procedure outlined in the text.

2.28 The charge conjugate spinor is $\psi_C = i\gamma^2\psi^*$. Show that $\bar{\psi}_C\psi_C = -\bar{\psi}\psi$. Show that the TCP product flips the electron spin: $iu^{(2)*}(p) = i\gamma^0\gamma^3\gamma^1 \times u^{(1)}(p)$, and that the same operation works for the positron spinors.

References

[1] F. Halzen and Alan D. Martin, *Quarks and Leptons* (John Wiley and Sons, New York, 1984).
[2] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1965).
[3] M. Peskin and D. Schroeder, *Introduction to Quantum Field Theory* (Addison Wesley, New York, 1995).
[4] V. Barger and Roger Phillips, *Collider Physics* (Addison Wesley, New York, 1987).
[5] J. J. Sakurai, *Advanced Quantum Mechanics* (Addison Wesley, New York, 1967).
[6] M. Veltman, *Diagrammatica* (Cambridge University Press, New York, 1994).
[7] B. Holstein, arXiv:hep-ph/0607187v1, 17 Jul 2006.

This page intentionally left blank

Chapter 3

A Little Field Theory

3.1. Introduction

We learned in Chapter 2 that Dirac's single electron wave equation when applied to the hydrogen atom described correctly the $2P^{3/2} - 2P^{1/2}$ fine structure splitting, but could not explain the $2P^{1/2} - 2S^{1/2}$ splitting — the Lamb shift, an effect that is 10% of the fine structure. The 1057 MHz level difference for $2P^{1/2} - 2S^{1/2}$ is called a radiative correction, and can be calculated by quantum electrodynamics, a quantum field theory. So, to understand nature, we need quantum field theory. The union of quantum mechanics and special relativity is quantum field theory. Much of the material of this chapter has been adopted from the excellent texts by Otto Nachtmann [1], and by Peskin and Schroeder [2]. By happy coincidence both of these texts use our metric.

There are classical fields. Maxwell's theory of electromagnetism is an example, where the $\vec{E}$ and $\vec{B}$ fields are physical objects — not just calculational aids — and causality is a fundamental component. The fields transmit information at a finite velocity, the speed of light. Lorentz transformations were originally derived to describe the effect of motion on Maxwell's equations. Quantum field theory treats Lorentz invariance and causality as basic premises. The well-established phenomenon of quantum entanglement in the Schrödinger theory has to be reconciled with causality. Maxwell's theory is very successful, but it has no photons, no matter how hard you look, until the field equations are quantized. Quantization of the electromagnetic field involves non-Hermitian creation and destruction operators that are very similar to raising and lowering operators of the harmonic oscillator. So, we begin with a review of the harmonic oscillator.

3.2. Harmonic Oscillator

The quantum oscillator can be solved from the Schrödinger equation or by algebraic techniques. Compared to solving the differential equation, the algebra makes it seem that the solution comes out of nowhere. The Hamiltonian is

$$H = \omega(P^2 + Q^2)/2, \tag{3.1}$$

$$P = p/\sqrt{m\omega}, \tag{3.2}$$

$$Q = x\sqrt{k/\omega}, \tag{3.3}$$

$$\omega^2 = k/m, \tag{3.4}$$

$$[x, p] = [Q, P] = i, \tag{3.5}$$

where Q and P are Herimitan operators. The raising and lowering operators are

$$a = (Q + iP)/\sqrt{2}, \tag{3.6}$$

$$a^\dagger = (Q - iP)/\sqrt{2}, \tag{3.7}$$

$$aa^\dagger - a^\dagger a = [a, a^\dagger] = 1, \tag{3.8}$$

$$H = \omega(aa^\dagger + a^\dagger a)/2, \tag{3.9}$$

$$\left[H, a^\dagger\right] = \omega a^\dagger, \tag{3.10}$$

$$[H, a] = -\omega a. \tag{3.11}$$

Equations (3.10) and (3.11) define $a^\dagger$ and a as raising and lowering operators, respectively. Solutions are kets $|n\rangle$, where n is an integer $0 \le n \le \infty$, and the matrix elements are

$$a\,|n\rangle = \sqrt{n}\,|n-1\rangle\,, \tag{3.12}$$

$$a\,|0\rangle = 0, \tag{3.13}$$

$$a^\dagger |n\rangle = \sqrt{n+1}\, |n+1\rangle\,, \tag{3.14}$$

$$H |n\rangle = \omega(n + 1/2)\, |n\rangle\,. \tag{3.15}$$

Remember that $\hbar = 1$. The constant term in the energy eigenvalue is called the zero point energy, and will be a problem in quantum field theory. The raising operator $a^\dagger$ takes the state up one level, adding energy ω. In field theory this is interpreted as creating a new particle with energy ω.

3.3. Lagrangian Density

In classical mechanics, one writes Lagrange's equations in terms of generalized coordinates and velocities:

$$\frac{d}{dt}\frac{\partial L(q^i, \dot{q}^i)}{\partial \dot{q}^j} - \frac{\partial L(q^i, \dot{q}^i)}{\partial q^j} = 0, \tag{3.16}$$

$$p^j = \frac{\partial L(q^i, \dot{q}^i)}{\partial \dot{q}^j}, \tag{3.17}$$

$$H(p^i, q^i) = \sum_j p^j \dot{q}^j - L(q^i, \dot{q}^i). \tag{3.18}$$

Usually $L = T - V$, where T is the kinetic energy, and V is the potential energy, and $\sum_i p^i \dot{q}^i = 2T$, so $H = T + V$. Lagrange's equations can be derived from a minimum principle, similar to the one proposed by Fermat to describe the reflection and refraction of light rays. Some people attribute an almost mystical significance to a minimum principle, as if nature should always choose the most efficient way. In the end, however, the minimum principle gives you a differential equation that you must solve. A very non-rigorous derivation is as follows:

$$\delta \int_a^b L(q^i, \dot{q}^i) dt = 0, \tag{3.19}$$

$$= \int_a^b ((\partial L/\partial q^i)\delta q^i + (\partial L/\partial \dot{q}^i)\delta \dot{q}^i) dt. \tag{3.20}$$

Integrate the second term by parts, and dispose of the constant term, because the variation is assumed to vanish at the end points:

$$\int_a^b (\partial L/\partial \dot{q}^i)\delta(dq^i/dt)dt = (\partial L/\partial \dot{q}^i)\delta q^i|_a^b - \int_a^b (d(\partial L/\partial \dot{q}^i)/dt)\delta q^i dt. \tag{3.21}$$

The variational time integral now contains an arbitrary common factor δq^i:

$$\int_a^b \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \right) \delta q^i dt = 0. \tag{3.22}$$

Lagrange's equations appear inside the parenthesis of the integral. The Lagrangian has units of energy, so the time integral has units (energy) × (time), like Planck's constant. It is sometimes called an action integral.

Field theory is Lagrangian based, but modifications to the particle version are required. First, in the non-relativistic particle theory, the position $\vec{x}$ is an observable, like the momentum $\vec{p}$, and they obey the commutation rule $[x, p_x] = i$. The time t is a parameter supplied by a clock outside the experiment. To be relativistically correct, space and time must be treated the same, and to do this the space coordinates $\vec{x}$ become parameters just like the time. The coordinates do not determine where the field is — the field is everywhere. The field amplitude $\phi(x)$ is a function of the coordinates, but they are no longer observables, and $\dot{x}$ has no meaning. In a sense it is the field amplitude itself that replaces the coordinates. This explains the appearance of functional derivatives in the Lagrange density for a scalar field. The Lagrangian in the integral of Equation (3.19) is replaced by a Lagrange density in order to add an integral over space d^3x, and obtain an invariant form d^4x. We introduce a shorthand for the derivative $\partial/\partial x^\mu = \partial_\mu$ in order to make the functional derivatives less cumbersome. For a classical scalar field one would write the action S:

$$S = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi). \tag{3.23}$$

The field and its derivatives are the degrees of freedom of the Lagrangian density. The variational principle proceeds exactly as above, except the integration by parts is four-dimensional. The result is

$$\delta S = \int d^4x \left(-\frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} + \frac{\partial \mathcal{L}}{\partial \phi} \right) \delta\phi = 0. \tag{3.24}$$

Note how the field and its four-dimensional derivatives have replaced the generalized coordinates and their time derivatives of Equation (3.22). Since $\delta\phi$ is arbitrary, the equations for the Lagrange density, an object that is often called the Lagrangian, are

$$\frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \tag{3.25}$$

The Lagrange density for the Klein–Gordon equation is:

$$\mathcal{L}(\phi, \partial_\mu \phi) = (\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2)/2, \tag{3.26}$$

$$\partial \mathcal{L}/\partial \phi = -m^2 \phi, \tag{3.27}$$

$$\partial \mathcal{L}/\partial(\partial_\mu \phi) = \partial^\mu \phi, \tag{3.28}$$

$$\frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = \partial_\mu \partial^\mu \phi + m^2 \phi = 0. \tag{3.29}$$

To obtain the third line, convert $(\partial_\mu \phi \partial^\mu \phi)/2 = (g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi)/2$. Note that the Lagrange density is relativistically invariant. The addition of an interaction term $-\lambda\phi^4/4$ to the Lagrange density gives the field equation:

$$(\partial^\mu \partial_\mu + m^2 + \lambda\phi^2)\phi = 0, \tag{3.30}$$

which, if the field has become the coordinates, looks like a harmonic oscillator! The term $\lambda\phi^2$ is a scalar, like the mass. Potentials in Lagrange densities that are functions of the field amplitude will be important later in our discussion of spontaneous symmetry breaking.

The Klein–Gordon Lagrangian can be written in terms of space and time derivatives as

$$\mathcal{L} = (\dot{\phi}^2 - (\nabla\phi)^2 - m^2\phi^2)/2, \tag{3.31}$$

$$\Pi = \partial\mathcal{L}/\partial(\dot{\phi}) = \dot{\phi}(x), \tag{3.32}$$

$$\mathcal{H} = \Pi(x)\dot{\phi}(x) - \mathcal{L} = (\dot{\phi}^2 + (\nabla\phi)^2 + m^2\phi^2)/2. \tag{3.33}$$

The middle equation is the canonical momentum, and Equation (3.33) is the Hamiltonian density. Note that the Hamiltonian density is not relativistically invariant (wrong sign on the gradient term).

3.4. Quantization of the Field

The real scalar field solution to the Klein–Gordon equation can be expanded in a Fourier series:

$$\phi(x) = (1/(2\pi)^3)\int d^3k/(2\omega)\big\{a(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega t)} + a^*(\vec{k})e^{-i(\vec{k}\cdot\vec{x}-\omega t)}\big\}. \tag{3.34}$$

$\phi(x)$ satisfies the Klein–Gordon equation if $\omega^2 - \vec{k}^2 - m^2 = 0$. $a(\vec{k})$ is a Fourier amplitude, and $d^3k/(2\omega)$ is Lorentz invariant as shown in Chapter 2. The factor $1/(2\pi)^3$ is chosen for convenience. The phase is the product of two four vectors $p^\mu x_\mu$. A similar formula was introduced for a Dirac spinor in Chapter 2, in order to calculate the expectation value of the velocity operator $\vec{\alpha}$.

We write the quantized field in capital letters: $\Phi(x)$, still a scalar, and still satisfying the Klein–Gordon equation, but now an operator, and an observable. A fundamental requirement of the observable operator is that two fields must commute if they cannot be causally connected:

$$[\Phi(x), \Phi(y)] = 0; (x-y)^2 < 0. \tag{3.35}$$

The fields are Lorentz invariant, and can be evaluated in any frame. There is an inertial frame where the two events are simultaneous, giving the equal time commutator:

$$[\Phi(\vec{x},t),\Phi(\vec{y},t)] = 0; \vec{x} \neq \vec{y}. \tag{3.36}$$

Of course the commutator also vanishes if $\vec{x} = \vec{y}$, because the field commutes with itself. So, we may drop the inequality, and simply write:

$$[\Phi(\vec{x},t),\Phi(\vec{y},t)] = 0. \tag{3.37}$$

How about $[\Phi(\vec{x},t),\dot{\Phi}(\vec{y},t)]$? Here things get interesting, because $\dot{\Phi}$ is the canonical momentum, and Φ, you remember, now acts like a coordinate, and $[x,p_x] = i$. If one writes $\dot{\Phi}(\vec{y},t) = (\Phi(\vec{y},t+dt) - \Phi(\vec{y},t))/dt$ it is clear that $[\Phi(\vec{x},t),\dot{\Phi}(\vec{y},t)] = 0$ if $\vec{x}$ and $\vec{y}$ differ by more than an infinitesimal. The correct answer is

$$[\Phi(\vec{x},t),\dot{\Phi}(\vec{y},t)] = i\delta(\vec{x}-\vec{y}), \tag{3.38}$$

where

$$\delta(\vec{x}-\vec{y}) = \frac{1}{(2\pi)^3}\int d^3k e^{i\vec{k}\cdot(\vec{x}-\vec{y})}. \tag{3.39}$$

This form of the Dirac three-dimensional delta function has already been encountered in Chapter 2, and occurs frequently in quantum field theory. If the phase in the integral is not zero, the integral vanishes because there are as many negative terms as positive ones. If the phase is zero, the integral diverges. The factor $(1/2\pi)^3$ normalizes the delta function. Look at one dimension:

$$\lim_{l\to\infty}\int_{-l}^{+l} dk e^{ikx} = (2\pi)\delta(x); \tag{3.40}$$

$$\int dx\delta(x) = (1/\pi)\int dx(\sin(xl))/x \tag{3.41}$$

$$= (1/\pi)\int_{-\infty}^{+\infty} dy(\sin y)/y = 1. \tag{3.42}$$

Here $y = xl$, and $\int_{-\infty}^{+\infty} dy(\sin y)/y = \pi$ is a standard integral. This explains the factor $(1/2\pi)^3$ that also appears on the Fourier expansion of $\phi(x)$.

The Fourier expansion of the quantized field has the same form as $\phi(x)$ in Equation (3.34). If $\phi(x)$ is real, then $\Phi(x)$ is Hermitian. The Fourier amplitudes $a(\vec{k})$ and $a^*(\vec{k})$ become operators, called destruction operator $a(\vec{k})$ and creation operator $a^\dagger(\vec{k})$, respectively. So, we write:

$$\begin{aligned}\Phi(\vec{x},t) &= (1/(2\pi)^3)\int d^3k/(2\omega)\{a(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega t)} \\ &\quad + a^\dagger(\vec{k})e^{-i(\vec{k}\cdot\vec{x}-\omega t)}\}, \qquad (3.43)\\ \dot{\Phi}(\vec{y},t) &= (i/(2\pi)^3)\int d^3k'/(2)\{-a(\vec{k'})e^{i(\vec{k'}\cdot\vec{y}-\omega' t)} \\ &\quad + a^\dagger(\vec{k'})e^{-i(\vec{k'}\cdot\vec{y}-\omega' t)}\}, \qquad (3.44)\\ [\Phi(\vec{x},t),\dot{\Phi}(\vec{y},t)] &= i\delta(\vec{x}-\vec{y}),\ \text{if} \qquad (3.45)\\ [a(\vec{k}),a(\vec{k'})] &= [a^\dagger(\vec{k}),a^\dagger(\vec{k'})] = 0, \qquad (3.46)\\ [a(\vec{k}),a^\dagger(\vec{k'})] &= (2\pi)^3 2\omega\delta(\vec{k}-\vec{k'}). \qquad (3.47)\end{aligned}$$

The last commutation rule is the analog of Equation (3.8) for the one-dimensional harmonic oscillator. The expression $(2\pi)^3 2\omega\delta(\vec{k}-\vec{k'})$ is a feature of our continuum normalization. The factor 2ω defines the normalization to be Lorentz invariant. In the box normalized scheme mentioned briefly in Chapter 2, $[a_{\vec{k}}a^\dagger_{\vec{k'}}] = \delta_{\vec{k},\vec{k'}}$, where the Kronecker δ is understood to mean $n_x = n'_x, n_y = n'_y, n_z = n'_z$, in complete agreement with Equation (3.8).

The only operators in the commutator are $a(\vec{k})$ and $a^\dagger(\vec{k})$. The commutators $[a(\vec{k}),a(\vec{k'})]$ and $[a^\dagger(\vec{k}),a^\dagger(\vec{k'})]$ must vanish to kill time dependent factors $e^{\pm i(\omega+\omega')t}$. The commutator $[a(\vec{k}),a^\dagger(\vec{k'})] = (2\pi)^3 2\omega\delta(\vec{k}-\vec{k'})$ gets rid of one $\int d^3k'$, and the time dependence is killed by $\omega = \omega'$, leaving us with the equation:

$$\begin{aligned}[\Phi(\vec{x},t),\dot{\Phi}(\vec{y},t)] &= (i/2)(1/(2\pi)^3)\int d^3k\{e^{i(\vec{k}\cdot(\vec{x}-\vec{y})} + e^{-i(\vec{k}\cdot(\vec{x}-\vec{y})}\} \\ &= i\delta(\vec{x}-\vec{y}). \qquad (3.48)\end{aligned}$$

The Hamiltonian is the volume integral of Equation (3.33):

$$H = \int d^3x\mathcal{H}. \qquad (3.49)$$

The operator $\mathcal{H}$ is a function of the field $\Phi(x)$ and its derivatives:

$$\mathcal{H} = (\dot{\Phi}^2 + (\nabla\Phi)^2 + m^2\Phi^2)/2, \tag{3.50}$$

$$\begin{aligned}\int d^3x\dot{\Phi}^2 &= -1/(2\pi)^3 \textstyle\int d^3k/4\{a(\vec{k})a(-\vec{k})e^{2i\omega t} + a^\dagger(\vec{k})a^\dagger(-\vec{k})e^{-2i\omega t}\\ &\quad - a(\vec{k})a^\dagger(\vec{k}) - a^\dagger(\vec{k})a(\vec{k})\},\end{aligned} \tag{3.51}$$

$$\begin{aligned}&\int d^3x(\nabla\Phi)^2\\ &\quad = 1/(2\pi)^3 \int d^3k/(2\omega)^2\vec{k}^2\{a(\vec{k})a(-\vec{k})e^{2i\omega t} + a^\dagger(\vec{k})a^\dagger(-\vec{k})e^{-2i\omega t}\\ &\qquad + a(\vec{k})a^\dagger(\vec{k}) + a^\dagger(\vec{k})a(\vec{k})\},\end{aligned} \tag{3.52}$$

$$\begin{aligned}&m^2\int d^3x\Phi^2\\ &\quad = m^2/(2\pi)^3 \int d^3k/(2\omega)^2\{a(\vec{k})a(-\vec{k})e^{2i\omega t} + a^\dagger(\vec{k})a^\dagger(-\vec{k})e^{-2i\omega t}\\ &\qquad + a(\vec{k})a^\dagger(\vec{k}) + a^\dagger(\vec{k})a(\vec{k})\},\end{aligned} \tag{3.53}$$

$$H = (1/(2\pi))^3 \int d^3k(a(\vec{k})a^\dagger(\vec{k}) + a^\dagger(\vec{k})a(\vec{k}))/4. \tag{3.54}$$

Equation (3.54) resembles Equation (3.9) for the harmonic oscillator. The three components are shown above, after integration over all space to give $\delta(\vec{k} + \vec{k}')$ for the time dependent terms, and $\delta(\vec{k} - \vec{k}')$ for the time independent ones. The sign of $\vec{k} \cdot \vec{k}'$ flips for the $\vec{k} = -\vec{k}'$ part. The time dependent sum vanishes because $-\omega^2 + \vec{k}^2 + m^2 = 0$. For the time independent terms the sign of ω^2 flips, to give $2\omega^2$ in the numerator. Analogs of Equations (3.10) and (3.11) are

$$[H, a(\vec{k}')] = -\omega' a(\vec{k}'), \tag{3.55}$$

$$[H, a^\dagger(\vec{k}')] = \omega' a^\dagger(\vec{k}'). \tag{3.56}$$

These operations are interpreted as subtracting or adding a particle of mass m and total energy $\omega' = \sqrt{\vec{k}'^2 + m^2}$ to the field. In the occupation number representation the ket $|n\rangle$ of the harmonic oscillator is replaced by a ket $|n_1, n_2, n_3, \ldots\rangle$ where n_i is the number of particles with momentum vectors $\vec{k_i}$ and energies ω_i. Of particular interest is the vacuum state $|0\rangle$ that has no particles, no energy, nothing. It is normalized to one: $\langle 0|0\rangle = 1$. By definition:

$$a(\vec{k})\,|0\rangle = 0, \tag{3.57}$$

$$a^{\dagger}(\vec{k})\,|0\rangle = |n_k = 1\rangle\,, \tag{3.58}$$

where the ket $|n_k = 1\rangle$ is not normalized (see Equation (3.61)). The vacuum expectation value of the Hamiltonian is

$$\langle 0|\,H\,|0\rangle = (1/(2\pi))^3 \int d^3k\,\langle 0|\,a(\vec{k})a^{\dagger}(\vec{k})\,|0\rangle\,/4 = \int d^3k\omega\delta^3(0)/2. \tag{3.59}$$

Here $\delta^3(0) = (1/2\pi)^3 \int d^3x$, which certainly looks divergent, but we fix it with a normalization volume $V = \int d^3x$, so that we have the vacuum expectation value:

$$\langle 0|\,H\,|0\rangle = V/(2\pi)^3 \int d^3k\omega(\vec{k})/2. \tag{3.60}$$

This is the field zero point energy — like the factor 1/2 in $E_n = \omega(n + 1/2)$. It is still divergent. This divergence is a common problem in field theory. To fix it, we subtract it off from the vacuum expectation value, asserting that it is an unobservable constant; we are interested in energy differences, and we set the vacuum energy to zero for convenience. The vacuum expectation value of the commutator is the same as our unwanted term:

$$\langle 0|\,a(\vec{k})a^{\dagger}(\vec{k}')\,|0\rangle = \langle 0|\,[a(\vec{k}), a^{\dagger}(\vec{k}')]\,|0\rangle = \langle\vec{k}|\vec{k}'\rangle = (2\pi)^3 2\omega\delta(\vec{k} - \vec{k}'). \tag{3.61}$$

The vacuum expectation value of the operator product $a(\vec{k})a^{\dagger}(\vec{k}')$ *is* the commutator, because the delta function is not an operator, and $\langle 0|0\rangle = 1$. The state vector normalization is characteristic of the continuum. We just subtract the commutator from the Hamiltonian:

$$H = (1/(2\pi))^3 \int d^3k(a(\vec{k})a^{\dagger}(\vec{k}) + a^{\dagger}(\vec{k})a(\vec{k}) - [a(\vec{k}), a^{\dagger}(\vec{k})])/4, \tag{3.62}$$

$$H = (1/(2\pi))^3 \int d^3k(a^{\dagger}(\vec{k})a(\vec{k}))/2, \tag{3.63}$$

$$\langle 0|\,H\,|0\rangle = 0. \tag{3.64}$$

Always lead with a destruction operator in order to have a finite vacuum expectation value. This is called 'normal ordering', and is denoted by a pair of colons. For boson fields, where the operators commute, the order can be shuffled without any sign changes. All observable quantities — energy, momentum, charge, current density, etc. for particles in the field — must have zero vacuum expectation value, and therefore be normal ordered. The operator product in a field expression can be displayed without the exponentials and integrals. For example, Φ^2 that appears in the Hamiltonian density has the operators:

$$\Phi^2 \sim (a_1 + a_1^\dagger)(a_2 + a_2^\dagger) = a_1 a_2 + a_1^\dagger a_2 + a_1 a_2^\dagger + a_1^\dagger a_2^\dagger. \qquad (3.65)$$

Normal ordering gives:

$$:\Phi^2: \sim a_1 a_2 + a_1^\dagger a_2 + a_2^\dagger a_1 + a_1^\dagger a_2^\dagger. \qquad (3.66)$$

The third term has been switched around. Each term now has zero vacuum expectation value.

Perhaps a word about quantum mechanical state vectors in field theory is appropriate. The occupation number kets are state vectors in momentum space, and we use continuum normalization, but the analog in configuration space, $\psi(\vec{x}, t)$ in Schrödinger's theory, that could describe a particle localized in space, does not exist. Schiff considers a non-relativistic particle of mass m confined in space to a Gaussian packet with width Δ centered at the origin at $t = 0$ [3]. Even in the absence of any external forces, the probability of finding the particle far away from the origin increases with time. The Gaussian width grows according to the formula:

$$\Delta(t)^2 = \Delta^2 + \frac{\hbar^2 t^2}{4m^2(\Delta^2)}. \qquad (3.67)$$

Now suppose we confine the particle initially to a size comparable to its Compton wave length: $\Delta = \lambda_c/(2\pi) = \hbar/(mc)$. Then the square root of the time dependent term in the formula above is $ct/2$! The tail of the wave function is growing at the speed of light. It turns out not to be possible to fix this in a relativistic theory, so there are no localized relativistic space wave functions with $\Delta \sim \lambda_c$ for material particles. The Compton wave length for a photon is not defined. Some reference to these problems for a free Dirac particle was made already in Chapter 2. The phenomena of wave function collapse and quantum entanglement are also in the realm of non-relativistic quantum mechanics. It is interesting that most of the experimental tests of quantum entanglement have used photons to carry the correlated information, and there is clearly no such thing as a non-relativistic photon. The theory does not include a localizable space wave function for a photon. A photon can be detected only by absorption or scattering. One might think that Maxwell's equations, that give light waves, and Schrödinger's equation, that gives matter waves, are on an equal footing, but this is deceptive. There are no photons without Planck's constant, and there are no matter waves without Planck's constant either. So, as $\hbar \to 0$, you have light waves and electron particles.

Let us call $|\vec{k}'\rangle$ the ket with occupation number $n(\vec{k}') = 1$. Then:

$$|\vec{k}'\rangle = a^\dagger(\vec{k}')\left|0\right\rangle, \tag{3.68}$$

$$H\left|\vec{k}'\right\rangle = (1/(2\pi))^3 \int d^3k(a^\dagger(\vec{k})a(\vec{k})a^\dagger(\vec{k}')\left|0\right\rangle/2 = \omega'|\vec{k}'\rangle, \tag{3.69}$$

$$\frac{\langle\vec{k}'|H|\vec{k}'\rangle}{\langle\vec{k}'||\vec{k}'\rangle} = \omega'. \tag{3.70}$$

We have used $a^\dagger(\vec{k})a(\vec{k})a^\dagger(\vec{k}')\left|0\right\rangle = (2\pi)^3 2\omega'\delta(\vec{k}-\vec{k}')a^\dagger(\vec{k})\left|0\right\rangle$. The matrix element $\langle\vec{k}'|H|\vec{k}'\rangle$ contains the vacuum expectation value of a product of four field operators:

$$\langle\vec{k}'|H|\vec{k}'\rangle = (1/(2\pi))^3 \int d^3k(\left\langle 0\right| a(\vec{k}')a^\dagger(\vec{k})a(\vec{k})a^\dagger(\vec{k}')\left|0\right\rangle)/2, \tag{3.71}$$

where we have used $(a^\dagger(\vec{k}')\left|0\right\rangle)^\dagger = \left\langle 0\right| a(\vec{k}')$. Vacuum expectation values of products of field operators occur frequently in quantum field theory. The vacuum expectation value of the product of an odd number of field operators vanishes, and the product must contain an a on the left, an $a^\dagger$ on the right, and the same number of a's and $a^\dagger$'s in between to be non-zero. The vacuum expectation value of a pair of operators is called a 'contraction', and a product of an even number of operators can be evaluated as a sum over all possible contractions. This is called Wick's theorem. The product in Equation (3.71) is typical in that it couples sharp states $\vec{k'}$ to continuum states $\vec{k}$. The contracted operators are linked by brackets over (or under) them. In this case there is only one combination that is not zero:

$$\langle 0|\, a' a^\dagger a a'^\dagger \,|0\rangle = \overbracket{a' a^\dagger}\,\overbracket{a a'}{}^\dagger, \tag{3.72}$$

$$\langle \vec{k'}|H|\vec{k'}\rangle = \omega' \langle \vec{k'}|\vec{k'}\rangle. \tag{3.73}$$

This technique can be used to evaluate products of harmonic oscillator operators (see Problem 3.10), and to prove Feynman's rules for evaluating diagrams.

Since the $a^\dagger$ commute among themselves, the ket with two particles is even under exchange:

$$|\vec{k_1}, \vec{k_2}\rangle = |\vec{k_2}, \vec{k_1}\rangle, \tag{3.74}$$

and the particles in our theory obey Bose–Einstein statistics. There are no forces so far in our theory, and the creation operators do not conserve energy. Something has to give up energy to create a particle with mass m and energy ω out of the vacuum.

The classical field is real, which means as we have seen in Chapter 2 that the probability density ρ and the current density $\vec{j}$ both vanish identically. Obviously this won't work for a charged field, so a complex field ϕ may be written in terms of two real fields ϕ_1 and ϕ_2: $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$, with ϕ having one sign of the charge, and ϕ^* having the other. The field operator $\Phi(x)$ is no longer Hermitian. All four of these fields have the same mass. The first example of this effect was in Chapter 1, Equations (1.23) and (1.24), where ϕ and ϕ^* are K^0 and $\bar{K}^0$, and ϕ_1 and ϕ_2 are K_S and K_L, respectively. The conserved quantity that distinguishes K^0 from $\bar{K}^0$ is strangeness, or hypercharge, rather than electric charge. The tiny mass difference between K_S and K_L, about $10^{-5}\,\text{eV}$, is an example of symmetry breaking, in this case non-conservation of hypercharge by the weak force.

The Fourier expansion for a charged field cross couples the creation and destruction operators:

$$\Phi(x) = (1/(2\pi))^3 \int d^3k/(2\omega)\big\{a(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega t)} + b^\dagger(\vec{k})e^{-i(\vec{k}\cdot\vec{x}-\omega t)}\big\}. \tag{3.75}$$

Here $a(\vec{k})$ destroys a positive particle, and $b^\dagger(\vec{k})$ creates a negative one. The operators change roles in the Hermitian conjugate field $\Phi^\dagger(x)$. Single operators for a charged field do not conserve charge, so the force that has not been specified has to ensure that the particles are created in pairs. Our field theory is incomplete.

3.5. Classical Electromagnetic Field

We begin by repeating Maxwell's equations from Chapter 1, in mks units:

$$\vec{\nabla}\cdot\vec{E} = \rho/\epsilon_0, \tag{3.76}$$

$$\vec{\nabla}\times\vec{E} = -\partial\vec{B}/\partial t, \tag{3.77}$$

$$\vec{\nabla}\cdot\vec{B} = 0, \tag{3.78}$$

$$\vec{\nabla}\times\vec{B} = \mu_0\vec{j} + (1/c^2)\partial\vec{E}/\partial t, \tag{3.79}$$

$$\vec{F} = q(\vec{E} + \vec{v}\times\vec{B}). \tag{3.80}$$

Take the curl of Equation (3.79):

$$\vec{\nabla}\times\vec{\nabla}\times\vec{B} = \mu_0\vec{\nabla}\times\vec{j} - (1/c^2)\partial^2\vec{B}/\partial t^2, \tag{3.81}$$

where we have used Equation (3.77) to eliminate the electric field. Now use the vector identity easily verified in Cartesian coordinates $\vec{\nabla}\times\vec{\nabla}\times = \vec{\nabla}(\vec{\nabla}\cdot) - \vec{\nabla}^2$, and $\vec{\nabla}\cdot\vec{B} = 0$ to obtain the wave equation with a source:

$$\vec{\nabla}^2\vec{B} - (1/c^2)\partial^2\vec{B}/\partial t^2 = -\mu_0\vec{\nabla}\times\vec{j}. \tag{3.82}$$

Taking the curl of Equation (3.77) and performing similar operations gives the source equation for the electric field:

$$\vec{\nabla}^2\vec{E} - (1/c^2)\partial^2\vec{E}/\partial t^2 = \mu_0\partial\vec{j}/\partial t + \vec{\nabla}\rho/\epsilon_0. \tag{3.83}$$

In the absence of sources, both fields satisfy the wave equation in free space, an equation for each of the three components that looks like the Klein–Gordon equation for $m = 0$.

$$\vec{\nabla}^2\vec{B} - (1/c^2)\partial^2\vec{B}/\partial t^2 = 0, \tag{3.84}$$

$$\vec{\nabla}^2\vec{E} - (1/c^2)\partial^2\vec{E}/\partial t^2 = 0. \tag{3.85}$$

The electromagnetic wave propagates in the vacuum with the speed of light in the direction of the wave number $\vec{k}$, where $|\vec{k}| = \omega/c$, and the electric and magnetic fields are perpendicular to each other, and mutually perpendicular to $\vec{k}$:

$$\vec{E}\cdot\vec{B} = \vec{E}\cdot\vec{k} = \vec{B}\cdot\vec{k} = 0. \tag{3.86}$$

The direction of polarization of the light wave is the electric vector. Solutions in the form of Fourier series, like the scalar field $\phi(x)$, may be written for both the electric and magnetic vectors:

$$\vec{E}(\vec{x},t) = (1/(2\pi))^3 \int d^3k/(2\omega)\big\{\vec{a}(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega t)} + \vec{a}^*(\vec{k})e^{-i(\vec{k}\cdot\vec{x}-\omega t)}\big\}, \tag{3.87}$$

$$\vec{B}(\vec{x},t) = (1/(2\pi))^3 \int d^3k/(2\omega)\big\{\vec{b}(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega t)} + \vec{b}^*(\vec{k})e^{-i(\vec{k}\cdot\vec{x}-\omega t)}\big\}. \tag{3.88}$$

The electric and magnetic fields are real. The vector $\vec{a}(\vec{k})$ points in the direction of the polarization. The transverse nature of the electromagnetic wave is specified by Maxwell's equations. In free space $\vec{\nabla}\cdot\vec{E} = 0$, so $\vec{k}\cdot\vec{a}(\vec{k}) = 0$, and Equation (3.79) gives the relation $\vec{k}\times\vec{a}(\vec{k}) = i\omega\vec{b}(\vec{k})$. It is worth emphasizing the basic nature of these relations, because we will be working not with the fields, but with the vector potential $\vec{A}$, where the transverse requirement is not automatic.

We have already met the vector potential. Take Equation (3.79), substitute $\vec{B} = \vec{\nabla}\times\vec{A}$ and $\vec{E} = -\vec{\nabla}V - \partial\vec{A}/\partial t$ to obtain:

$$\vec{\nabla}\times\vec{\nabla}\times\vec{A} = \mu_0\vec{j} + 1/c^2(-\vec{\nabla}\partial V/\partial t - \partial^2\vec{A}/\partial t^2). \tag{3.89}$$

Making the vector identity substitution and re-arranging terms gives a form of the wave equation for the vector potential:

$$\vec{\nabla}(\vec{\nabla} \cdot \vec{A} + (1/c^2)\partial V/\partial t) - \vec{\nabla}^2 \vec{A} + (1/c^2)\partial^2 \vec{A}/\partial t^2 = \mu_0 \vec{j}. \quad (3.90)$$

We have the usual wave equation for $\vec{A}$ provided that:

$$\vec{\nabla} \cdot \vec{A} + (1/c^2)\partial V/\partial t = 0. \quad (3.91)$$

Note that the vector potential, unlike the magnetic field, is parallel to the current density. Equation (3.91) is a gauge condition, called the Lorentz gauge. This choice of gauge also gives a wave equation for V, derived from Equation (3.76):

$$\vec{\nabla}^2 V - (1/c^2)\partial^2 V/\partial t^2 = \rho/\epsilon_0. \quad (3.92)$$

If V is a constant, then $\vec{\nabla}\cdot\vec{A} = 0$, and $\vec{A}$ is parallel to $\vec{E}$, and satisfies the transversality condition $\vec{k}\cdot\vec{A} = 0$. This is called the radiation gauge, and is the one commonly used in quantum electrodynamics, although it is not a Lorentz invariant requirement. In the radiation gauge we may write the Fourier expansion in the same form as for the electric vector, where the amplitude $\vec{\alpha}(\vec{k})$ points in the direction of the polarization of the light:

$$\vec{A}(\vec{x}, t) = (1/(2\pi))^3 \int d^3k/(2\omega)\{\vec{\alpha}(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega t)} + \vec{\alpha^*}(\vec{k})e^{-i(\vec{k}\cdot\vec{x}-\omega t)}\}. \quad (3.93)$$

The Lorentz gauge condition does not uniquely constrain the vector potential, since it is unaffected by a gauge transformation of the form $\vec{A'} = \vec{A} + \vec{\nabla}\chi$, where χ satisfies the wave equation: $\nabla^2\chi - (1/c^2)\partial^2\chi/\partial t^2 = 0$.

3.6. Four-Dimensional Form

Current and charge densities combine to form a four vector, as do the vector and scalar potentials:

$$J^\mu = (c\rho, \vec{j}), \quad (3.94)$$

$$A^\mu = (V/c, \vec{A}), \quad (3.95)$$

$$\partial_\mu A^\mu = 0, \tag{3.96}$$

$$\partial^\mu \partial_\mu A^\nu = \mu_0 J^\nu \tag{3.97}$$

$$A'^\mu = A^\mu + \partial^\mu \chi, \tag{3.98}$$

$$\mu_0 \epsilon_0 = (1/c^2). \tag{3.99}$$

Equation (3.96) is the Lorentz gauge condition. The components of Equation (3.97) are the wave equations (3.90) and (3.92), and Equation (3.98) is a gauge transformation. These formulas are all in mks units. Remember that $\partial_0 = \partial/\partial ct$. We shall soon go over to natural units, where all of the constants ϵ_0, μ_0, c are set to one.

There are three components of the electric field (E_x, E_y, E_z), and three components of the magnetic field (B_x, B_y, B_z), and they mix in a Lorentz transformation. They fit comfortably into a four-dimensional six-vector, a totally antisymmetric second rank tensor that is the analog of a curl called $F^{\mu\nu}$:

$$\begin{aligned} F^{\mu\nu} &= \partial^\mu A^\nu - \partial^\nu A^\mu \\ &= \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{pmatrix}. \end{aligned} \tag{3.100}$$

The space–space terms, like F^{12} are components of $\vec{B} = \vec{\nabla} \times \vec{A}$. The lower component tensor $F_{\mu\nu}$ has the signs of the electric field reversed, because there is only one space component that changes sign, but the magnetic field, which has two space components, stays the same. The transposed tensor $F^{\nu\mu}$ reverses all of the signs. The free field Lagrange density is

$$\mathcal{L} = -F_{\mu\nu}F^{\mu\nu}/4 = -(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu)/4, \tag{3.101}$$

$$= (\vec{E}^2/(c^2) - \vec{B}^2)/2. \tag{3.102}$$

The Lagrange density is a function only of the derivatives $\partial_\mu A_\nu$ and $\partial^\mu A^\nu$, and does not depend on the field A_μ, so $\partial\mathcal{L}/\partial A_\mu = 0$, meaning no mass term in the wave equation. Lagrange's equations become:

$$\frac{\partial}{\partial x^\mu}\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = \partial_\mu(-\partial^\mu A^\nu + \partial^\nu A^\mu) = -\Box A^\nu + \partial^\nu(\partial_\mu A^\mu) = 0. \tag{3.103}$$

The Lorentz gauge condition eliminates the second term, leaving the $m = 0$ wave equation for the vector potential:

$$\Box A^\mu = \partial_\nu \partial^\nu A^\mu = 0. \tag{3.104}$$

The four vector A^μ has been reduced to only two components in order to describe electromagnetic plane waves. The time component $A^0 = V/c$ has been set to zero, and the space vector $\vec{A}$ is perpendicular to the wave number $\vec{k}$. These two conditions are neither Lorentz nor gauge invariant. If $\vec{\epsilon}$ is the polarization vector, then $\vec{\epsilon} \cdot \vec{k} = 0$ is not Lorentz invariant; $\epsilon^\mu k_\mu = 0$ is Lorentz invariant, but what is ϵ^0? We shall return to this problem directly. The Fourier expansion of Equation (3.92) has the same form as the electric field, and the polarization is in the same direction.

The canonical four momentum is defined by

$$\Pi^\alpha = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\alpha)} = -\frac{\partial A^\alpha}{c\partial t} + \partial^\alpha A^0. \tag{3.105}$$

The time component vanishes: $\Pi^0 = 0$. The space components are the electric field:

$$\vec{\Pi} = \vec{E}/c. \tag{3.106}$$

The Hamiltonian density is

$$\mathcal{H} = \Pi^\mu \partial A_\mu/(\partial ct) - \mathcal{L} = -\vec{E} \cdot \dot{\vec{A}}/c^2 - (\vec{E}^2/(c^2) - \vec{B}^2)/2. \tag{3.107}$$

Equation (3.107) doesn't look right. $\mathcal{H} = (\vec{E}^2/(c^2) + \vec{B}^2)/2$ is the correct form. To fix it, we may add the four divergence of any function that vanishes at infinity to the Hamiltonian density. By Gauss' theorem the integral equals the function evaluated on the surface, where it vanishes as the surface becomes large. Another way of saying the same thing is that the Lagrangian and Hamiltonian densities are defined up to the addition of an arbitrary four divergence of a function that vanishes at infinity. Adding the term $\partial_\mu(A_0 F^{0\mu})$ to $\mathcal{H}$ gives the correct expression.

3.7. Quantization of the Electromagnetic Field

For the quantized field we return to natural units, so in Maxwell's equations $\epsilon_0 = \mu_0 = c = 1$. The factor of $1/(4\pi)$ remains in Coulomb's law, so that the fine structure constant is $\alpha = e^2/(4\pi) = 1/(137.036)$. The four-vector potential is quantized as a Hermitian operator:

$$\partial^\mu A_\mu = 0, \tag{3.108}$$

$$\Box A_\mu = 0, \tag{3.109}$$

$$A_\mu(x) = (1/(2\pi)^3) \int d^3k/(2\omega)\{a_\mu(\vec{k})e^{-ik_\nu x^\nu} + a^\dagger_\mu(\vec{k})e^{ik_\nu x^\nu}\}, \tag{3.110}$$

$$k^\nu k_\nu = \omega^2 - \vec{k}^2 = 0. \tag{3.111}$$

Equation (3.108) is the Lorentz gauge condition, (3.109) is the wave equation — the Klein–Gordon equation for $m = 0$ — and in the Fourier expansion the operators satisfy the requirement of the Lorentz gauge:

$$k^\nu a_\nu(\vec{k}) = k^\nu a^\dagger_\nu(\vec{k}) = 0. \tag{3.112}$$

The commutation rules for the operators are as follows:

$$[a_\mu(\vec{k}), a_\nu(\vec{k'})] = [a^\dagger_\mu(\vec{k}), a^\dagger_\nu(\vec{k'})] = 0, \tag{3.113}$$

$$[a_\mu(\vec{k}), a^\dagger_\nu(\vec{k'})] = -g_{\mu\nu}(2\pi)^3 2\omega\delta(\vec{k} - \vec{k'}). \tag{3.114}$$

The metric tensor guarantees that, for the physically significant space components a_j the commutator is positive. Apart from the four-dimensional index, meaning that different components of a and $a^\dagger$ commute, these requirements are the same as we used for the scalar field. Unfortunately, the Lorentz gauge condition is not consistent with the commutation rules, since:

$$k^\mu a_\mu = 0 \text{ means } k_\nu (2\pi)^3 2\omega\delta(\vec{k} - \vec{k'}) = 0, \tag{3.115}$$

and that makes no sense. It turns out not to be possible to reconcile the gauge condition with the operator form of the vector potential, but it can be satisfied as an expectation value for a physical photon state. The argument due to Gupta and Bleuler is given by Nachtmann on page 104 of his text, mentioned at the beginning of this chapter. See Problem 3.12.

We introduce state vectors in analogy with the scalar field. The vacuum state $|0\rangle$ has no photons; by definition $a_\mu(\vec{k})\,|0\rangle = 0$; and a photon with wave number $\vec{k}$ and four-dimensional polarization vector ϵ^μ is given by

$$|\vec{k}, \epsilon\rangle = -\epsilon^\mu a^\dagger_\mu(\vec{k})\,|0\rangle\,. \tag{3.116}$$

For a real photon with wave number $\vec{k}$ the polarization has two components, both perpendicular to $\vec{k}$, and the time component $\epsilon_0 = 0$. Therefore, the four momentum product $\epsilon_\mu k^\mu = 0$, and $\vec{\epsilon} \cdot \vec{k} = 0$. The first dot product is Lorentz invariant, but the second one is not unless the transformation is collinear with the photon direction. Consider a Lorentz transformation to a frame moving with velocity v along the z axis, with $\vec{k}$ and $\vec{\epsilon}$ in the (x, z) plane:

$$k_z = k\cos\theta, \tag{3.117}$$

$$k_x = k\sin\theta, \tag{3.118}$$

$$k_0 = |\vec{k}| = \omega, \tag{3.119}$$

$$\epsilon_z = -\sin\theta, \tag{3.120}$$

$$\epsilon_x = \cos\theta, \tag{3.121}$$

$$\epsilon_0 = 0, \tag{3.122}$$

$$\omega' = \gamma\omega(1 + v\cos\theta), \tag{3.123}$$

$$\epsilon'_0 = -\gamma v\sin\theta, \tag{3.124}$$

$$k'_\mu \epsilon'^\mu = k_\mu \epsilon^\mu = 0, \tag{3.125}$$

$$\vec{k'} \cdot \vec{\epsilon'} = \omega' \epsilon'_0 = -\gamma\omega' v\sin\theta. \tag{3.126}$$

Now add to ϵ'_μ a part parallel to k'_μ: $\epsilon'_\mu \to \epsilon'_\mu + ak'_\mu$ such that $\epsilon'_0 \to 0$. The constant $a = -\epsilon'_0/\omega' = \gamma v \sin\theta/\omega'$, and now $\vec{\epsilon'} \cdot \vec{k'} = 0$. This works because $k_\mu k^\mu = 0$, and would not work for a massive particle. So, it is always possible to retain transverse waves after a Lorentz transformation, but it does not happen automatically. Plane waves can be described in the (x, y) plane for $\vec{k} = \omega\hat{z}$: $\epsilon_x = \cos\phi$, and $\epsilon_y = \sin\phi$. Circular polarization states are $\epsilon^+ = -(\epsilon_x + i\epsilon_y)/\sqrt{2}$ and $\epsilon^- = (\epsilon_x - i\epsilon_y)/\sqrt{2}$, where the sign of ϵ^+ is chosen to agree with the spherical harmonic Y_1^{+1}.

The calculation of the Hamiltonian proceeds like the scalar case, apart for the vector polarization. The Hamiltonian density is

$$\mathcal{H} = (\vec{E}^2 + \vec{B}^2)/2 = ((\partial\vec{A}/\partial t)^2 + (\vec{\nabla} \times \vec{A})^2)/2, \tag{3.127}$$

where the vector potential has two space vector polarization components $\vec{\epsilon}^{\,1}$, and $\vec{\epsilon}^{\,2}$ with Fourier amplitudes $a^{(s)}(\vec{k})$ and $a^{\dagger(s)}(\vec{k})$, where $s = 1, 2$, satisfying the conditions $\vec{\epsilon}^{(s)} \cdot \vec{k} = 0$, and $\vec{\epsilon}^{\,(r)} \cdot \vec{\epsilon}^{\,(s)} = \delta_{r,s}$. Thus,

$$\vec{A}(\vec{x}, t) = (1/(2\pi))^3 \int d^3k/(2\omega) \sum_{s=1}^{2} \{\vec{\epsilon}^{\,s} a^{(s)}(\vec{k}) e^{i(\vec{k}\cdot\vec{x}-\omega t)} + \vec{\epsilon}^{\,*s} a^{\dagger(s)}(\vec{k}) e^{-i(\vec{k}\cdot\vec{x}-\omega t)}\}. \tag{3.128}$$

The commutation rules are

$$[a^{(r)}(\vec{k}), a^{(s)}(\vec{k'})] = 0, \tag{3.129}$$

$$[a^{\dagger(r)}(\vec{k}), a^{\dagger(s)}(\vec{k'})] = 0, \tag{3.130}$$

$$[a^{(r)}(\vec{k}), a^{\dagger(s)}(\vec{k'})] = (2\pi)^3(2\omega)\delta_{r,s}\delta(\vec{k} - \vec{k'}). \tag{3.131}$$

The calculation of the electric field term is straightforward:

$$\begin{aligned}\int d^3x(\partial\vec{A}/\partial t)^2 \\ &= -(1/(2\pi))^3 \int d^3k/4 \sum_{s=1}^{2} \{a^{(s)}(\vec{k})a^{(s)}(-\vec{k})e^{-2i\omega t} \\ &\quad + a^{\dagger(s)}(\vec{k})a^{\dagger(s)}(-\vec{k})e^{2i\omega t} - a^{(s)}(\vec{k})a^{\dagger(s)}(\vec{k}) - a^{\dagger(s)}(\vec{k})a^{(s)}(\vec{k})\}.\end{aligned} \tag{3.132}$$

The magnetic field involves $\vec{\nabla} \times \vec{A}$, or terms like $\vec{k} \times \vec{\epsilon}^{\,s}$ in the Fourier expansion. These terms may be reduced by noting that $\vec{k} \times \vec{\epsilon}^{\,1} = \omega\vec{\epsilon}^{\,2}$, and $\vec{k} \times \vec{\epsilon}^{\,2} = -\omega\vec{\epsilon}^{\,1}$. As in the scalar case, terms in $\int d^3x\vec{B}^2$ that involve $\delta(\vec{k} + \vec{k}')$ have flipped signs, so that the time dependent terms cancel, leaving the Hamiltonian:

$$\int d^3x\mathcal{H} = (1/(2\pi))^3 \int d^3k/4 \sum_s \{a^{(s)}(\vec{k})a^{\dagger(s)}(\vec{k}) + a^{\dagger(s)}(\vec{k})a^{(s)}(\vec{k})\}. \tag{3.133}$$

This formula looks very much like Equation (3.54) for the scalar field, and it has the same divergence problem for the vacuum expectation value $\langle 0| H |0\rangle$. We use the same fix: subtract the commutator to kill the $aa^\dagger$ part. This is equivalent to carrying out normal ordering (signified by the colons) that reverses the order on the first term, doubling the second one, and leaving us with:

$$H = \int d^3x : \mathcal{H} := (1/(2\pi))^3 \int d^3k/2 \sum_s a^{\dagger(s)}(\vec{k})a^{(s)}(\vec{k}). \tag{3.134}$$

Now $\langle 0| H |0\rangle = 0$, and

$$\begin{aligned} &\langle \vec{k}', \epsilon'^1 | H | \vec{k}', \epsilon'^1 \rangle \\ &\quad = (1/(2\pi))^3 \int d^3k/2 \sum_s \langle 0| a^{(1)}(\vec{k}')a^{\dagger(s)}(\vec{k})a^{(s)}(\vec{k})a^{\dagger(1)}(\vec{k}') |0\rangle . \end{aligned} \tag{3.135}$$

The vacuum expectation value involves the product $a'a^\dagger aa'^\dagger$, coupling sharp and continuum states in pairs. The result, using the commutation rules, is

$$\langle \vec{k}', \epsilon'^1 | H | \vec{k}', \epsilon'^1 \rangle = \omega' \langle \vec{k}', \epsilon'^1 | \vec{k}', \epsilon'^1 \rangle. \tag{3.136}$$

'Normal ordering' means always leading with a destruction operator, so that for any observable quantity M, $\langle 0| M |0\rangle = 0$.

3.8. Cross-Section Calculations in Schrödinger's Quantum Mechanics

Golden Rule #2 for the calculation of a transition rate in perturbation theory was introduced in Chapter 2 in the context of three-body phase space. It is useful to derive this formula, a procedure introduced by Dirac, before we confront similar formulas in the relativistic theory. For one thing, it points out how different are the worlds of non-relativistic and relativistic quantum mechanics. Non-relativistic theory is built for bound states, with time independent Hamiltonians that have infinite sets of discreet eigenvalues and corresponding eigenfunctions. Scattering problems can be treated, but bound states are the foundation. Perturbation theory gives a technique for making small corrections to bound state solutions. The quantum theory of fields, on the other hand, is, at least for our purposes, only a perturbation theory, and bound states are usually not obtained by perturbation expansions. The Schrödinger theory is used to describe the energy levels of charmonium $c\bar{c}$ and bottomonium $b\bar{b}$ bound states, where one can apply non-relativistic quantum mechanics with good results because of the heavy quark mass, using a potential that is inspired by quantum chromodynamics. Non-relativistic theory can be applied to bound states where the binding energy is small compared to the masses. This approximation does not work for the light quarks, and lattice gauge theory, mentioned in Chapter 1 in the context of quark masses, has been used instead. Lattice gauge theory attempts exact bound state solutions in field theory by working in discreet, rather than continuous space time, and is beyond the scope of this book. The time dependent Schrödinger equation $H\psi(\vec{x},t) = i\hbar\partial\psi(\vec{x},t)/\partial t$ can be written:

$$\psi(\vec{x},t) = e^{-iHt/\hbar}\psi(\vec{x},0). \tag{3.137}$$

Then the expectation value of an arbitrary observable Hermitian operator A can be written:

$$\langle A\rangle = \int d^3x\psi^*(\vec{x},0)e^{iHt/\hbar}Ae^{-iHt/\hbar}\psi(\vec{x},0). \tag{3.138}$$

If $[A,H] = 0$, A is a constant of the motion. If the time displacement operators $e^{\pm iHt/\hbar}$ are attached to the wave functions, then A is time independent, and one has what is called the 'Schrödinger picture'. On the other hand, one may take the wave functions to be constant, and attach the time displacement operators to the observable. That is called the 'Heisenberg picture'. It all depends on how you group the terms in the expectation value expression. A time displacement operator $U(t, t_0)$ for the state vector may be defined:

$$\psi(t) = U(t, t_0)\psi(t_0); U(t, t_0) = e^{-iH(t-t_0)/\hbar}. \tag{3.139}$$

If H is Hermitian $H^\dagger = H$, then U is unitary: $U^\dagger(t, t_0) = U^{-1}(t, t_0) = U(t_0, t)$. The 'interaction picture', originally introduced by Dirac, splits the time dependence into two parts, one from the unperturbed Hamiltonian, and the other from the perturbation term. The perturbation term switches on at $t = 0$, so for $t > 0$ the Hamiltonian consists of two parts:

$$(H_0 + H_I)\psi(\vec{x}, t) = i\hbar\partial\psi(\vec{x}, t)/\partial t, \tag{3.140}$$

and the wave function $\psi_I(\vec{x}, t)$ for $t > 0$ in the interaction representation is defined as

$$\psi_I(\vec{x}, t) = e^{iH_0 t/\hbar}\psi(\vec{x}, t). \tag{3.141}$$

The derivation of the golden rule starts with the assumption that the unperturbed Hamiltonian can be solved exactly:

$$H_0\psi_0(\vec{x}, t) = i\hbar\partial\psi_0(\vec{x}, t)/\partial t;\ H_0 u_n(\vec{x}) = E_n u_n(\vec{x}). \tag{3.142}$$

The general solution to the unperturbed wave equation may be written as an expansion in terms of $u_n(\vec{x})e^{-iE_n t/\hbar}$ with constant coefficients a_n:

$$\psi_0(\vec{x}, t) = \sum_n a_n u_n(\vec{x})e^{-iE_n t/\hbar}; \tag{3.143}$$

as may be verified by direct substitution. If the perturbation H_I is switched on at $t = 0$, we have plenty of time $-\infty < t < 0$ to define the initial state as one of the unperturbed states u_k, E_k. The complete wave function in Equation (3.140) $\psi(\vec{x}, t)$ can be expanded in the eigenfunctions of H_0, but now the coefficients are time dependent. This is the interaction picture mentioned above; the unperturbed time dependence is explicitly factored out:

$$\psi(\vec{x}, t) = \sum_n a_n(t)u_n(\vec{x})e^{-iE_n t}. \tag{3.144}$$

The interaction representation wave function defined by Equation (3.141) has no energy dependent phase factors:

$$\psi_I(\vec{x}, t) = \sum_n a_n(t) u_n(\vec{x}), \tag{3.145}$$

and a time displacement operator for the wave function in the interaction representation can be defined:

$$\psi_I(t) = U_I(t, t_0)\psi_I(t_0); U_I(t, t_0) = e^{iH_0 t/\hbar} U(t, t_0) e^{-iH_0 t_0/\hbar}, \tag{3.146}$$

where $U(t, t_0) = e^{-i(H_0+H_I)(t-t_0)/\hbar}$. Exponentials involving operators are tricky to handle, because the usual rules only apply if $[H_0, H_I] = 0$, and that is almost never true. For instance, a potential that varies in space does not commute with the momentum operator. However, some operations are straightforward. If we define the perturbation Hamiltonian in the interaction representation as

$$H_I(t) = e^{iH_0 t/\hbar} H_I e^{-iH_0 t/\hbar}, \tag{3.147}$$

then $i\hbar \partial\psi_I(t)/\partial t = H_I(t)\psi_I(t)$. The perturbation determines the time dependence of $\psi_I(t)$, and H_0 determines the time dependence of the operators. It is not difficult to show (no problems with who commutes with what) that:

$$\frac{i\hbar \partial U_I(t, t_0)}{\partial t} = H_I(t) U_I(t, t_0). \tag{3.148}$$

The solution to this first order differential equation with the boundary condition $U_I(t_0, t_0) = 1$ will lead us to a formalism for perturbation calculations to various orders in H_I. This is a convenient representation for quantum field theory calculations because, while the total Hamiltonian $H_0 + H_I$ is not Lorentz invariant, the interaction term, which is same as in the Lagrangian except for a change in algebraic sign, *is* Lorentz invariant. See Equation (3.204) for $H_I = -e(\bar{\psi}\gamma^\mu\psi)A_\mu$ in quantum electrodynamics, that is manifestly covariant (the dot product of two four vectors).

First we return to the traditional derivation of the first-order Golden rule. Substitution of Equation (3.144) into the wave equation gives the following equivalent form, so far with no approximations:

$$\dot{a}_m(t) = (1/i\hbar)\sum_n a_n(t)e^{-i(E_n-E_m)t/\hbar} \times \int d^3x(u_m^*(\vec{x})H_I u_n(\vec{x})), \tag{3.149}$$

where we have used the orthonormality of the eigenfunctons: $\int d^3xu_m^*(\vec{x})u_n(\vec{x}) = \delta_{m,n}$. Incidentally, if we were to absorb the energy phases into the matrix element for H_I we would obtain the matrix element for $H_I(t)$. Now we make an approximation. We assume the system is in the state k at $t = 0$, and that to first order in H_I, $a_k(0) = 1$, $a_n(0) = 0$, for $n \neq k$, and $\dot{a}_k(t) = 0$. Any depletion of the initial state is assumed to be second order. Then for any state m the wave equation to first order becomes:

$$\dot{a}_m(t) = (1/i\hbar)a_k(0)e^{-i(E_k-E_m)t/\hbar} \times \langle m|\, H_I \,|k\rangle\,. \tag{3.150}$$

Defining $\omega_{km} = (E_k - E_m)/\hbar$ and integrating from 0 to t, with H_I independent of time in that range, and $a_k(0) = 1$, gives:

$$a_m(t) = (2/(i\hbar))\,\langle m|\, H_I \,|k\rangle\, e^{-i\omega_{km}t/2}\sin(\omega_{km}t/2)/\omega_{km}. \tag{3.151}$$

The probability of a transition $k \to m$ as a function of time is

$$|a_m(t)|^2 = (4/\hbar^2)|\,\langle m|\, H_I \,|k\rangle\,|^2\sin^2(\omega_{km}t/2)/\omega_{km}^2. \tag{3.152}$$

The function $\sin^2(\omega_{km}t/2)/\omega_{km}^2$ has the form of the single slit diffraction pattern in optics: $\sin^2(x)/x^2$. For $\omega_{km} = 0$ the central peak value is $t^2/4$, and the first zero is at $\omega_{km}t/2 = \pi$, so the area under the central peak in ω space is proportional to the time. We are interested in time proportional transitions that conserve energy $\omega_{km} \approx 0$, so we multiply by a density of final states $dN/d\omega$ and integrate over $d\omega$ (dN is dimensionless):

$$\begin{aligned}&\int d\omega|a_m(t)|^2dN/d\omega\\ &\quad = (4/\hbar^2)|\,\langle m|\, H_I \,|k\rangle\,|^2\int d\omega dN/d\omega\sin^2(\omega_{km}t/2)/\omega_{km}^2.\end{aligned} \tag{3.153}$$

Treating the strong maximum at $\omega = 0$ as 'δ function like', we may factor out $dN/d\omega|_{\omega=0}$. Then $\int d\omega\sin^2(\omega t/2)/\omega^2 = (\pi t/2)$, giving us the transition rate formula:

$$w_{k\to m} = (2\pi/\hbar^2)|\langle m| H_I |k\rangle|^2 dN/d\omega = (2\pi/\hbar)|\langle m| H_I |k\rangle|^2 dN/dE. \tag{3.154}$$

A variation may be written in terms of initial state i and final state f, with an energy δ function, and multiplication by the density of final states dN/dE and integration over energy understood:

$$w_{i\to f} = (2\pi/\hbar)|\langle f| H_I |i\rangle|^2 \delta(E_i - E_f). \tag{3.155}$$

This is the celebrated Golden Rule #2. The derivation may seem like humble beginnings, but its applications are far and wide.

3.9. Rutherford Scattering

As an example of the application of the golden rule to a scattering problem, we will solve the Coulomb scattering of an α particle (no spin) by a very heavy nucleus. The nucleus does not recoil, so the α particle energy is conserved $|\vec{k_i}| = |\vec{k_f}| = k$, but the direction of the momentum vector changes: $\vec{k_i} - \vec{k_f} = \vec{K}$, where the momentum transfer $K = 2k\sin(\theta/2)$, θ being the polar scattering angle: $\vec{k_i}\cdot\vec{k_f} = k^2\cos(\theta)$. We will keep Planck's constant $\hbar$ in the calculation throughout, to show that it cancels out of Rutherford's formula. Rutherford's original derivation used classical mechanics, with no $\hbar$'s anywhere to be seen, and it is one of those happy accidents that the quantum mechanical treatment gives the same answer. Although in our field theory work we use continuum normalization, we will here use box normalization, as an exercise in a different technique. Quantum field theory can also be done using the box normalization, replacing integrals with sums. This approach has already been introduced in Chapter 2.

So imagine a large cubic box of dimension L with the origin in the center. Each coordinate runs $-L/2 \le x \le +L/2$. The wave number components are quantized by having periodic boundary conditions: $e^{-ik_xL/2} = e^{+ik_xL/2}$. This condition gives continuous wave motion for boxes placed nose to tail, and results in wave numbers that are proportional to integers: $k_x = (2\pi n_x/L)$, and similarly for k_y and k_z.

Treating the large quantum numbers as quasi-continuous gives phase space for the final state:

$$\begin{aligned} dN/dE &= (dn_x dn_y dn_z)/dE \\ &= (L/(2\pi))^3 (dk_x dk_y dk_z)/dE = V/(2\pi)^3 d^3k/dE. \end{aligned} \tag{3.156}$$

The α particle is non-relativistic: $E = \hbar^2 k^2/(2m)$, $p = \hbar k$, and $d^3k = k^2 dk d\Omega$, where $d\Omega$ is the solid angle into which it is scattered. So, the phase space factor becomes:

$$dN/dE = V/(2\pi)^3 (mp/\hbar^3) d\Omega. \tag{3.157}$$

Now we have to evaluate the matrix element. The Coulomb potential energy is $V(r) = (ZZ'e^2)/r$, which we multiply by a convergence factor $e^{-\alpha r}$ that will be set $\alpha \to 0$ after integration. The matrix element then is

$$\langle f | H_I | i \rangle = (ZZ'e^2/V) \int d\Omega' \int r' dr' e^{(-\alpha + iK\cos\theta')r'}. \tag{3.158}$$

The volume V in the denominator comes from the box normalized wave functions $\psi_i = (1/\sqrt{V}) e^{i\vec{k_i}\cdot\vec{x}}$. The primed angles are integrated over (they are not the scattering angles!). The z' axis is oriented along the momentum transfer vector $\vec{K} = \vec{k_i} - \vec{k_f}$ to simplify the integration. The result is

$$\langle f | H_I | i \rangle = (ZZ'e^2/V) \times (4\pi/(K^2 + \alpha^2)). \tag{3.159}$$

The transition rate can be written as a differential per unit solid angle; after setting the convergence factor $\alpha \to 0$ we have:

$$\frac{dw_{i\to f}}{d\Omega} = \frac{m(ZZ'e^2)^2}{4Vp^3 \sin^4(\theta/2)}. \tag{3.160}$$

By definition the cross-section is the transition rate divided by the flux. There is one α particle in the box of volume V, giving a density $\rho = 1/V$; and with velocity $v = p/m$, the flux $= \rho v = v/V$, and the Rutherford cross-section is

$$\frac{d\sigma}{d\Omega} = \frac{(ZZ'e^2)^2}{16E^2 \sin^4(\theta/2)}, \tag{3.161}$$

where $E = p^2/(2m)$ is the α particle kinetic energy, and $p = mv = \hbar k$ is a constant in the scattering. The normalization volume cancels out, and Planck's constant has disappeared. This same result can be obtained from the Born approximation, which treats the scattered wave as a perturbation in the time independent theory. Having worked so hard to get rid of Planck's constant, we now stick it back in, because $\alpha = e^2/(\hbar c)$ is so convenient for numerical calculations. For 5 MeV α particles on gold, we have:

$$\frac{d\sigma}{d\Omega} = \frac{(ZZ')^2(\alpha\hbar c)^2}{16E^2\sin^4(\theta/2)} = \frac{(2\times 79)^2(197)^2\times 10^{-26}}{16\times 25\times(137)^2\sin^4(\theta//2)}\text{cm}^2, \quad (3.162)$$

where we have used $\hbar c = 197$ eV-nm converted to MeV-cm. The product of atomic numbers ZZ' almost cancels α. The arithmetic gives $d\sigma/d\Omega = 1.3\times 10^{-24}\,\text{cm}^2/\sin^4(\theta/2)$, or about 1 barn divided by $\sin^4(\theta/2)$. This is a large Coulomb cross-section, due to the high atomic number of the gold nucleus.

The total Coulomb cross-section $\int d\Omega(d\sigma/d\Omega)$ is not defined. The divergence of Rutherford's formula in the forward direction is a feature of the Coulomb potential, and is sometimes called the Coulomb pole. If we had retained the convergence factor α in the potential $V(r) = (ZZ'e^2)e^{-\alpha r}/r$, then the Fourier transform would not diverge at $\theta = 0$, and the total cross-section would be finite. In scattering from a gold atom there would be a natural cutoff of the Coulomb force due to the negative charge of the electron shells at very high impact parameters. In any event, $0 \leq \theta \leq \pi$ is a polar angle, so there are no more zeros in $\sin(\theta/2)$, and an integral over any range $\theta_0 \leq \theta \leq \pi$ is well defined, and in fact easy to integrate. For example the cross-section for the entire backward hemisphere is

$$(2\pi)\int_{\pi/2}^{\pi}\sin(\theta)d\theta(d\sigma/d\Omega) = 1.3\times 10^{-24}\,\text{cm}^2\times(4\pi). \quad (3.163)$$

3.10. Green's Function for Helmholtz's Equation

In perturbation theory we may be interested in a solution to an inhomogeneous wave equation of the form:

$$(H - E_k)\psi(x) = f(x). \quad (3.164)$$

An often occurring variation on this problem, since the wave equation is homogeneous, is an interaction term of the form $V(x)\psi(x)$ that gives an integral equation for the wave function. But for now we assume that $f(x)$ is a known function, and that there are continuous eigenfunction and eigenvalue solutions to the homogeneous equation with $f(x) = 0$:

$$Hu_k(x) = E_k u_k(x), \tag{3.165}$$

$$\int dx u_k^*(x) u_{k'}(x) = \delta(k - k'), \tag{3.166}$$

$$\int dk u_k^*(x) u_k(x') = \delta(x - x'). \tag{3.167}$$

Equation (3.167) is often called 'completeness'. We can solve the inhomogeneous equation if we can find a Green's function:

$$(H - E_k) G_k(x, x') = \delta(x - x'), \tag{3.168}$$

then

$$\psi(x) = \int dx' G_k(x, x') f(x'), \tag{3.169}$$

this solves the problem up to an arbitrary solution to the homogeneous equation that may be required by the boundary conditions. The Green's function can be written as a product of eigenfunctions of H:

$$G_k(x, x') = \int dk' \frac{u_{k'}^*(x') u_{k'}(x)}{E_{k'} - E_k}, \tag{3.170}$$

because H only operates on the eigenfunction $u_{k'}(x)$, the energy denominator cancels:

$$(H - E_k) G_k(x, x') = \int dk' u_{k'}^*(x') u_{k'}(x) = \delta(x - x'). \tag{3.171}$$

For a free particle in three dimensions $-\nabla^2$ replaces H, and E_k becomes k^2, with $u_{\vec{k}}(\vec{x}) = (1/2\pi)^{3/2} e^{i\vec{k}\cdot\vec{x}}$; and the Green's function becomes:

$$G_k(\vec{x}, \vec{x'}) = (1/(2\pi))^3 \int d^3k' \frac{e^{i\vec{k'}\cdot(\vec{x} - \vec{x'})}}{(k'^2 - k^2)}. \tag{3.172}$$

There are two zero's in the denominator at $k' = \pm k$. Choose the z axis in k' space along the vector $\vec{x} - \vec{x'}$ and carry out the angular integration to obtain

$$G_k(\vec{x}, \vec{x'}) = \frac{-i}{8\pi^2} \int_{-\infty}^{+\infty} \frac{k'dk'}{|\vec{x} - \vec{x'}|(k' - k)(k' + k)} \{e^{ik'|\vec{x}-\vec{x'}|} - e^{-ik'|\vec{x}-\vec{x'}|}\}. \tag{3.173}$$

The extra factor of two in the denominator comes from extending the even function from $-\infty \to +\infty$ on the k' axis. The integral in the complex k' plane is along the real axis, with poles at $k' = \pm k$. The first term $e^{ik'|\vec{x}-\vec{x'}|}$ converges in the upper half plane, where $\mathrm{Im}\{k'\} > 0$; and the second term $e^{-ik'|\vec{x}-\vec{x'}|}$ converges in the lower half plane where $\mathrm{Im}\{k'\} < 0$. Choose the contour so that the path circles the pole at $k' = +k$ for the upper half plane and $k' = -k$ in the lower half plane. In the first integral the pole is circled counter clockwise $(+2\pi i)$, and in the second clockwise $(-2\pi i)$, so that the two terms add, and the final result is

$$G_k(\vec{x}, \vec{x'}) = \frac{e^{ik|\vec{x}-\vec{x'}|}}{4\pi|\vec{x} - \vec{x'}|}. \tag{3.174}$$

Substituting $k \to i\alpha$ gives the damped Coulomb potential as the Greens function for the inhomogeneous Helmholtz equation:

$$(\nabla^2 - \alpha^2)G(\vec{x}, \vec{x'}) = \delta(\vec{x} - \vec{x'}); G(\vec{x}, \vec{x'}) = \frac{-e^{-\alpha|\vec{x}-\vec{x'}|}}{4\pi|\vec{x} - \vec{x'}|}. \tag{3.175}$$

The Fourier transform of this Green's function has already been calculated for the Rutherford cross-section:

$$G(\vec{k}, \vec{k'}) = \frac{-1}{(\vec{k} - \vec{k'})^2 + \alpha^2}. \tag{3.176}$$

The Green's function for Poisson's equation $\vec{\nabla}^2 G(\vec{x}, \vec{x'}) = \delta(\vec{x} - \vec{x'})$ is

$$G(\vec{x}, \vec{x'}) = -\frac{1}{(4\pi|\vec{x} - \vec{x'}|)}. \tag{3.177}$$

This function satisfies Laplace's equation everywhere except at the point $\vec{x} = \vec{x'}$. The singularity can be isolated by setting $\vec{x'} = 0$, integrating over a sphere of radius r_0, and using Gauss's theorem to convert to a surface integral:

$$\int d\vec{S} \cdot \vec{\nabla}(1/r) = -4\pi; \tag{3.178}$$

independent of the radius r_0. The same function in momentum space can be obtained by setting $\alpha = 0$ in Equation (3.176). The Green's function for Poisson's equation is familiar from electrostatics. Its Fourier transform $G(\vec{k}, \vec{k'}) = -1/(\vec{k} - \vec{k'})^2$ when cast in four-dimensional language, and adding the metric tensor to account for photon spin, will become the photon propagator in quantum electrodynamics.

3.11. Time Ordering

We return to the scalar field that satisfies the Klein–Gordon equation:

$$(\partial^2/\partial t^2 - \vec{\nabla}^2 + m^2)\Phi(\vec{x}, t) = 0, \tag{3.179}$$

$$\Phi(\vec{x}, t) = \Phi(x) = (1/(2\pi)^3) \int (d^3k/(2\omega))\{a(\vec{k})e^{-ik\cdot x} + a^\dagger(\vec{k})e^{ik\cdot x}\}, \tag{3.180}$$

$$[a(\vec{k}), a(\vec{k'})] = [a^\dagger(\vec{k}), a^\dagger(\vec{k'})] = 0, \tag{3.181}$$

$$[a(\vec{k}), a^\dagger(\vec{k'})] = (2\pi)^3 2\omega\delta(\vec{k} - \vec{k'}), \tag{3.182}$$

where $\omega^2 = \vec{k}^2 + m^2$. The commutator of the field at two different space–time points is a straightforward calculation. The only terms that do not vanish are $[a(\vec{k}), a^\dagger(\vec{k'})]$ and $[a^\dagger(\vec{k}), a(\vec{k'})]$.

$$[\Phi(x), \Phi(y)] = \frac{1}{(2\pi)^3} \int \frac{d^3k}{2\omega} \{e^{-ik\cdot(x-y)} - e^{ik\cdot(x-y)}\}. \tag{3.183}$$

If $(x - y)^2 < 0$, so that the distance in space between the two events is greater than can be covered by a light signal in the time difference $(x_0 - y_0)$, then the two fields must commute. This is the microcausality that we invoked at the beginning of this chapter. The commutator as written is Lorentz invariant, so it can be evaluated in any inertial frame, and the fact that $(x - y)^2 < 0$ means that there is a frame where $x_0 = y_0$, our equal time commutator. If the time dependent terms in the exponents vanish, we have:

$$[\Phi(x), \Phi(y)] = \frac{1}{(2\pi)^3} \int \frac{d^3k}{2\omega} \{e^{-i\vec{k}\cdot(\vec{x}-\vec{y})} - e^{i\vec{k}\cdot(\vec{x}-\vec{y})}\}, \tag{3.184}$$

and this expression vanishes, because it is an odd function of $\vec{k}$, and the integral is over all positive and negative values. So microcausality is preserved. If $(x-y)^2 > 0$, the two events can be correlated, and the commutator need not vanish. Since the commutator does not involve any field operators, it is equal to its vacuum expectation value:

$$\begin{aligned} \langle 0| [\Phi(x), \Phi(y)] |0\rangle &= [\Phi(x), \Phi(y)] \\ &= \frac{1}{(2\pi)^3} \int \frac{d^3k}{2\omega} \{e^{-ik\cdot(x-y)} - e^{ik\cdot(x-y)}\}. \end{aligned} \tag{3.185}$$

One term in the commutator, $\langle 0| \Phi(x)\Phi(y) |0\rangle$ is called a propagator, and written $D(x-y)$. The equal time condition is dropped:

$$D(x-y) = \langle 0| \Phi(x)\Phi(y) |0\rangle = \int \frac{d^3k}{2\omega} e^{-ik\cdot(x-y)}. \tag{3.186}$$

Obviously

$$\langle 0| [\Phi(x), \Phi y] |0\rangle = D(x-y) - D(y-x). \tag{3.187}$$

The formula for $D(x-y)$ can be written as a four-dimensional integral, with various prescriptions regarding the treatment of the singularity. The Feynman propagator for the scalar field that satisfies the Klein–Gordon equation is written:

$$D_F(x-y) = \int \frac{d^4k}{(2\pi)^4} \frac{ie^{-ik\cdot(x-y)}}{(k^2 - m^2 + i\epsilon)}. \tag{3.188}$$

Here, $d^4k = d^3k d\omega$, and $-\infty < \omega < +\infty$, that is the energy can take on negative as well as positive values. The denominator is $\omega^2 - (\vec{k}^2 + m^2) + i\epsilon$, where $\epsilon \to 0$ from positive values. Define $\omega_0^2 = \vec{k}^2 + m^2$. Then we may write the denominator as $(\omega - (\omega_0 - i\epsilon))(\omega + (\omega_0 - i\epsilon))$, where ϵ has been redefined, but no matter, since it is still ϵ. There are two poles in the complex ω plane, one at $\omega_0 - i\epsilon$, and one at

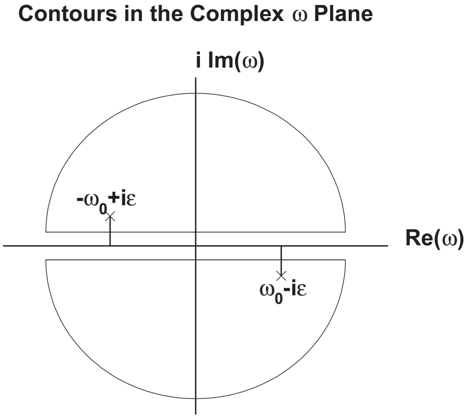


Figure 3.1: Complex ω plane for contour integration of time ordered products. If $x_0 > y_0$, then the integral on the semi-circle converges downward, and circles the pole at $\omega_0 - i\epsilon$. If $x_0 < y_0$, the integral converges upward, circling the pole at $-\omega_0 + i\epsilon$.

$-\omega_0 + i\epsilon$. The integral becomes:

$$D_F(x-y) = \int \frac{d^3k}{(2\pi)^4} e^{i\vec{k}\cdot(\vec{x}-\vec{y})} \int d\omega \frac{ie^{-i\omega(x_0-y_0)}}{(\omega-(\omega_0-i\epsilon))(\omega+(\omega_0-i\epsilon))}. \tag{3.189}$$

The integration is similar to the one performed for the non-relativistic Green's function. The contour along the real axis and the two poles are shown in Figure 3.1. The exponent is $e^{-i\omega(x_0-y_0)}$, so if $\mathrm{Im}\,\{\omega\} < 0$ (lower half plane) $x_0 > y_0$ for convergence, or y occurs before x, so that $\Phi(x)\Phi(y)$ is *time ordered*, then circle the pole at $+\omega_0$, with a factor of $-2\pi i$ for clockwise rotation, to get the equation:

$$x_0 > y_0$$

$$D_F(x-y) = \frac{(-2\pi i)i}{(2\pi^4)} \int d^3k e^{i\vec{k}\cdot(\vec{x}-\vec{y})} \frac{e^{-i\omega_0(x_0-y_0)}}{2\omega_0}, \tag{3.190}$$

$$D_F(x-y) = \frac{1}{(2\pi)^3} \int d^3k \frac{e^{-ik\cdot(x-y)}}{2\omega}, \tag{3.191}$$

where in the last line we have dropped the subscript on ω, the energy. The factor involving ω_0 in Equation (3.190) is the residue of the complex integration.

Now suppose that $y_0 > x_0$, that is x occurs first. Then for convergence on the infinite semicircle the contour must be closed in the upper half plane, circling the pole at $\omega = -\omega_0$ counter clockwise, and picking up a factor of $+2\pi i$. The sign changes in numerator and denominator cancel each other, but the phase changes sign ($\vec{k} \to -\vec{k}$ without penalty), so that x and y are interchanged:

$$\begin{aligned} &y_0 > x_0 \\ &D_F(x-y) = \frac{1}{(2\pi)^3}\int d^3k \frac{e^{ik\cdot(y-x)}}{2\omega}, \end{aligned} \tag{3.192}$$

which is the same as Equation (3.191). In other words, $D_F(x-y)$ is the vacuum expectation value of the time ordered product:

$$D_F(x-y) = \langle 0|\, T\Phi(x)\Phi(y)\, |0\rangle. \tag{3.193}$$

The formal definition of a time ordered product of two scalar fields is

$$T\Phi(x)\Phi(y) = \theta(x_0 - y_0)\Phi(x)\Phi(y) + \theta(y_0 - x_0)\Phi(y)\Phi(x), \tag{3.194}$$

where the step function $\theta(x) = 0,\ x < 0$ and $\theta(x) = 1,\ x > 0$. The derivative is a delta function: $d\theta(x)/dx = \delta(x)$. The four-dimensional integral Equation (3.188) is a Green's function for the Klein–Gordon equation:

$$(\partial^2/\partial x_0^2 - \nabla^2 + m^2)D_F(x-y) = -\int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot(x-y)} = -\delta^{(4)}(x-y), \tag{3.195}$$

since the numerator and denominator cancel out as $\epsilon \to 0$.

The four-dimensional Fourier transform of $D_F(x-y)$ can be calculated from Equation (3.188):

$$\begin{aligned} &\int d^4x D_F(x-y)e^{ik'\cdot(x-y)} \\ &\quad = \int d^4k \frac{ie^{i(k-k')\cdot y}}{k^2 - m^2 + i\epsilon}(1/(2\pi)^4)\int d^4x e^{i(k'-k)\cdot x} = \frac{i}{k'^2 - m^2 + i\epsilon}. \end{aligned} \tag{3.196}$$

It is this form of the propagator in momentum space that will appear in calculations of transition rates. The scalar formula is the simplest to work out, but will not be useful until the Higgs scalar in Chapter 8. We shall derive similar formulas for the photon and electron that we will use all of the time. Propagators describe exchanged particles in intermediate states. Four momentum is conserved at each vertex, so that the intermediate state particle is 'virtual', or not on mass shell, and the ϵ in the denominator can be set to zero without fear, because the expression, in the scalar case $(k^2 - m^2)$, does not vanish.

The photon propagator can be written down almost by inspection of the scalar expression, with the metric tensor, that appears in the photon commutation rules, taking care of the polarization. The mass is set to zero:

$$\langle 0| T A_\mu(x) A_\nu(y) |0\rangle = \int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot(x-y)} \left\{ \frac{-ig_{\mu\nu}}{(k^2 + i\epsilon)} \right\}. \tag{3.197}$$

The photon propagator in momentum space is displayed in the curly brackets, with $\epsilon \to 0$.

3.12. Quantization of the Dirac Field

The free Dirac field Lagrange density is written in terms of both ψ and $\bar{\psi}$, treated as independent fields:

$$\mathcal{L} = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x). \tag{3.198}$$

Then the usual procedure, with respect to ψ, gives:

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} = i\bar{\psi}\gamma^\mu, \tag{3.199}$$

$$\frac{\partial \mathcal{L}}{\partial \psi} = -m\bar{\psi}, \tag{3.200}$$

$$\frac{\partial}{\partial x^\mu} \partial\mathcal{L}/\partial(\partial_\mu \psi) - \frac{\partial \mathcal{L}}{\partial \psi} = i\frac{\partial \bar{\psi}}{\partial x^\mu}\gamma^\mu + m\bar{\psi} = 0. \tag{3.201}$$

Equation (3.201) is the correct adjoint Dirac equation. To obtain the Dirac equation for $\psi(x)$, write:

$$i\bar{\psi}\gamma^{\mu}\frac{\partial\psi}{\partial x^{\mu}} = i\frac{\partial(\bar{\psi}\gamma^{\mu}\psi)}{\partial x^{\mu}} - i\frac{\partial\bar{\psi}}{\partial x^{\mu}}\gamma^{\mu}\psi, \tag{3.202}$$

where the divergence term vanishes upon integration, or, what amounts to the same thing, integrate by parts. Then proceed as above to obtain the usual Dirac equation:

$$i\gamma^{\mu}\frac{\partial\psi}{\partial x^{\mu}} - m\psi = 0. \tag{3.203}$$

The Lagrange density for quantum electrodynamics is

$$\mathcal{L} = -F_{\mu\nu}F^{\mu\nu}/4 + \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi + eA_{\mu}\bar{\psi}\gamma^{\mu}\psi, \tag{3.204}$$

where the interaction term comes from the substitution $i\partial_{\mu} \to i\partial_{\mu} + eA_{\mu}$ in the free Dirac Lagrangian density. This is the four-dimensional equivalent of the substitution $\vec{p} \to \vec{p} - e\vec{A}$ in the non-relativistic Hamiltonian. The algebraic sign of the $\vec{A}$ term flips for the Hamiltonian. The elementary charge in this formula is the magnitude $e = |e|$. This substitution is gauge invariant, provided that the gauge transformation on the vector potential $A_{\mu} \to A_{\mu} + \partial_{\mu}\chi$ is accompanied by a phase change on the Dirac wave function: $\psi(x) \to \psi(x)e^{ie\chi(x)}$. Substituting these changes into the Lagrangian, and using $\bar{\psi} \to \bar{\psi}e^{-ie\chi(x)}$ leaves the expression invariant. $e\bar{\psi}\gamma_{\mu}\psi$ is the electron current density, and $\mathcal{L}$ couples the current to the vector potential. The interaction term and the free field term can be combined by defining the *covariant derivative*:

$$D_{\mu} = \partial_{\mu} - ieA_{\mu}; \text{then } \mathcal{L} = -F_{\mu\nu}F^{\mu\nu}/4 + \bar{\psi}(i\gamma^{\mu}(D_{\mu} - m)\psi. \tag{3.205}$$

The two terms in this Lagrangian density are manifestly gauge invariant, because for $D'_{\mu} = D_{\mu} - ie\partial_{\mu}\chi$:

$$D'_{\mu}e^{ie\chi}\psi = e^{ie\chi}D_{\mu}\psi, \tag{3.206}$$

and the exponential is canceled by $\bar{\psi}e^{-ie\chi}$. This is called a local $U(1)$ gauge transformation. A global one would have a constant χ. Gauge invariance is an interesting feature of Maxwell's equations that follows from the introduction of the vector potential as a calculational aid, but the same procedure can be used to generate the electromagnetic interaction in the Lagrange density. Start with the free field Dirac Lagrangian and the electromagnetic Lagrangian, the first two terms in Equation (3.204). Then make the gauge transformation $\psi(x) \rightarrow e^{ie\chi(x)}\psi(x)$, and identify the vector potential $A_\mu(x) = \partial_\mu \chi(x)$. Note that the coupling constant e appears explicitly in the phase. The gauge transformation treats the two fields differently. For the Dirac field it is a phase that is a function of four-dimensional space time, while for the vector field A_μ it adds a derivative. The mass term for spin 1/2, $m\bar{\psi}\psi$, is gauge invariant, but that is not true for the vector field. A term $-M^2 A_\mu A^\mu$ acquires an added piece $A_\mu \partial^\mu \chi + \partial_\mu \chi A^\mu + \partial_\mu \chi \partial^\mu \chi$. The vector field must be massless for the Lagrangian to be gauge invariant. The commutator of two D's is proportional to $F_{\mu\nu}$, and is not a derivative operator:

$$[D_\mu, D_\nu] = -ie([A_\mu, \partial_\nu] + [\partial_\mu, A_\nu]) = -ieF_{\mu\nu}. \tag{3.207}$$

This form is obtained because the two components of the vector potential commute $[A_\mu, A_\nu] = 0$. Quantum electrodynamics is called an *Abelian* gauge theory.

The Dirac field contains positive and negative energies, that we interpret as electrons and positrons, both with positive energy. It is a charged field, and like the charged scalar field, it must be complex, with creation and destruction operators that cross couple the particle–antiparticle pair. To achieve this we write:

$$\psi(x) = \frac{1}{(2\pi)^3} \int \frac{d^3p}{2E} \sum_{s=1,2} \left\{ a_s(\vec{p}) u_s(p) e^{-ip\cdot x} + b_s^\dagger(\vec{p}) v_s(p) e^{ip\cdot x} \right\}, \tag{3.208}$$

here $a_s(\vec{p})$ destroys an electron with spin s and momentum $\vec{p}$, and $b_s^\dagger(\vec{p})$ creates a positron of the same spin and momentum. These operators are scalars in the space of the γ matrices, while $u_s(p)$ and $v_s(p)$ are four component column vectors. The electron spinors $u_s(p)$ for

$s = 1, 2$ were defined in Chapter 2, Equations (2.116) and (2.117), while the positron spinors are given together in Equation (2.155). The quantized field $\psi(x)$ is a four component operator. The adjoint field operator $\bar{\psi} = \psi^\dagger \gamma^0$ is

$$\bar{\psi}(x) = \frac{1}{(2\pi)^3} \int \frac{d^3p}{2E} \sum_{s=1,2} \left\{ a_s^\dagger(\vec{p}) \bar{u}_s(p) e^{ip\cdot x} + b_s(\vec{p}) \bar{v}_s(p) e^{-ip\cdot x} \right\}. \tag{3.209}$$

Electrons and positrons obey the Pauli exclusion principle. It turns out that the correct way to guarantee the Pauli principle is to require the operators to *anticommute.* Defining the anticommutator:

$$\{A, B\} \equiv AB + BA, \tag{3.210}$$

we have the following rules:

$$\begin{aligned} \left\{a_s(\vec{p}), a_{s'}(\vec{p'})\right\} &= \left\{b_s(\vec{p}), b_{s'}(\vec{p'})\right\} \\ &= \left\{a_s^\dagger(\vec{p}), a_{s'}^\dagger(\vec{p'})\right\} = \left\{b_s^\dagger(\vec{p}), b_{s'}^\dagger(\vec{p'})\right\} = 0, \qquad (3.211) \\ \left\{a_s(\vec{p}), b_{s'}(\vec{p'})\right\} &= \left\{a_s^\dagger(\vec{p}), b_{s'}^\dagger(\vec{p'})\right\} \\ &= \left\{a_s(\vec{p}), b_{s'}^\dagger(\vec{p'})\right\} = \left\{a_s^\dagger(\vec{p}), b_{s'}(\vec{p'})\right\} = 0, \qquad (3.212) \end{aligned}$$

interchanging any two $a's$ or $b's$ gives a change of sign: $aa' = -a'a$, etc. In particular, $(a^\dagger)^2 = 0$, so you cannot create two electrons in the same state. There are two more anticommutation rules:

$$\left\{a_s(\vec{p}), a_{s'}^\dagger(\vec{p'})\right\} = \left\{b_s(\vec{p}), b_{s'}^\dagger(\vec{p'})\right\} = (2\pi)^3 2E \; \delta_{s,s'} \delta(\vec{p} - \vec{p'}). \tag{3.213}$$

The energy $E = \sqrt{p^2 + m^2}$, where electron and positron have precisely the same mass, is always positive. Anticommutation is a characteristic of spin 1/2. The Pauli spin matrices anticommute, as do the Dirac matrices γ^μ. The field operators are different in that their squares are zero, while $\sigma_j^2 = 1$, and $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$. Fermion field operators are an example of Grassmann numbers in mathematics. They anticommute, their squares are zero, and their derivatives are equal to integrals (rather like exponentials in this regard).

The Hamiltonian can be obtained from the Lagrangian by the usual procedure. The canonical momentum is

$$\Pi(x) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \psi)} = i\bar{\psi}(x)\gamma^0 = i\psi^\dagger(x). \tag{3.214}$$

Note that $\bar{\psi}$ plays a role similar to $\dot{\Phi}$ in the scalar field. Then the Hamiltonian density is

$$\mathcal{H} = \Pi \cdot \frac{\partial \psi}{\partial x_0} - \mathcal{L} = \bar{\psi}(-i\vec{\gamma} \cdot \vec{\nabla} + m)\psi;\ H = \int d^3x \mathcal{H}. \tag{3.215}$$

The time derivatives cancel out, and we are left with the space factor $-i\vec{\gamma} \cdot \vec{\nabla} + m$, that is the correct form for the Dirac Hamiltonian.

The anticommuting field operators deserve some practice for familiarity. First let us calculate the Hamiltonian, an operator that we have evaluated for the scalar and electromagnetic cases already. Write out the field operators:

$$\psi(x) = \frac{1}{(2\pi)^3} \int \frac{d^3p}{2E} \sum_{s=1,2} [a_s(\vec{p})u_s(p)e^{-ip\cdot x} + b_s^\dagger(\vec{p})v_s(p)e^{ip\cdot x}], \tag{3.216}$$

$$\bar{\psi}(x) = \frac{1}{(2\pi)^3} \int \frac{d^3p'}{2E'} \sum_{s'=1,2} [a_{s'}^\dagger(\vec{p'})\bar{u}_{s'}(p')e^{ip'\cdot x} + b_{s'}(\vec{p'})\bar{v}_{s'}(p')e^{-ip'\cdot x}]. \tag{3.217}$$

Then note that the operator in the middle of $\mathcal{H}$, $(-i\vec{\gamma} \cdot \vec{\nabla} + m)$ becomes:

$$(\vec{\gamma} \cdot \vec{p} + m)u_s(p) = \gamma^0 E u_s(p) \tag{3.218}$$

on the first term, and

$$(-\vec{\gamma} \cdot \vec{p} + m)v_s(p) = -\gamma^0 E v_s(p) \tag{3.219}$$

on the second term. These equations follow from $(\not{p} - m)u_s(p) = 0$ and $(\not{p} + m)v_s(p) = 0$, respectively. The factor γ^0 changes $\bar{u}$ and $\bar{v}$ back into $u^\dagger$ and $v^\dagger$, and after the space integration produces delta functions $\delta(\vec{p} - \vec{p'})$ the orthogonality and normalization come in handy:

$$u^\dagger_{s'}(p)u_s(p) = v^\dagger_{s'}(p)v_s(p) = 2E\delta_{ss'}, \tag{3.220}$$

and for terms with $\delta(\vec{p}+\vec{p}')$:

$$u^\dagger_{s'}(\vec{p})v_s(-\vec{p}) = 0. \tag{3.221}$$

The result is

$$H = \frac{1}{(2\pi)^3}\frac{1}{2}\sum_s \int d^3p(a^\dagger_s(\vec{p})a_s(\vec{p}) - b_s(\vec{p})b^\dagger_s(\vec{p})), \tag{3.222}$$

which after normal ordering becomes:

$$H = \frac{1}{(2\pi)^3}\frac{1}{2}\sum_s \int d^3p(a^\dagger_s(\vec{p})a_s(\vec{p}) + b^\dagger_s(\vec{p})b_s(\vec{p})). \tag{3.223}$$

Note the sign change on the second term when the order of the operators is reversed. Also note that the sign is the same for electrons and positrons — they both have positive energy. As usual, we define the vacuum state $|0\rangle$ such that $a_s(\vec{p})\,|0\rangle = 0$, and $a^\dagger_s(\vec{p})\,|0\rangle$ is a state with one electron with momentum $\vec{p}$ and spin s. Normal ordering has guaranteed that $\langle 0|\,H\,|0\rangle = 0$. We may calculate the single state expectation value using the anticommutation rule Equation (3.213):

$$\langle 0|\,a_{s'}(\vec{p'})Ha^\dagger_{s'}(\vec{p'})\,|0\rangle = E'\,\langle 0|\,a_{s'}(\vec{p'})a^\dagger_{s'}(\vec{p'})\,|0\rangle. \tag{3.224}$$

The energies of the particles are the eigenvalues of the Hamiltonian. The exclusion principle is satisfied, and the state vector is odd under the exchange of any two electrons. Electrons and positrons can be distinguished, because they have opposite charges, but among the positrons the exclusion principle is in full force.

The analog of the equal time commutator $[\Phi(\vec{x},t),\dot{\Phi}(\vec{y},t)] = i\delta(\vec{x}-\vec{y})$ for the scalar field is the anticommutator $\left\{\psi(x),\bar{\psi}(y)\right\}$. This is an unusual quantity, in that it is a matrix written one way, and a number written the other, so we will work it out in component form. The indices α and β refer to the columns in $u_s(p)$ and the rows in $\bar{u}_s(p)$ respectively in the space of the γ matrices:

$$\begin{aligned}
&\{\psi_\alpha(x), \bar{\psi}_\beta(y)\} \\
&\quad = \frac{1}{(2\pi)^6} \sum_{s,s'} \int \frac{d^3p}{2E} \int \frac{d^3p'}{2E'} [\{a_s(\vec{p}), a^\dagger_{s'}(\vec{p'})\} \\
&\qquad \times (u_s(p))_\alpha (\bar{u}_{s'}(p'))_\beta e^{-i(p\cdot x - p'\cdot y)} \\
&\qquad + \{b^\dagger_s(\vec{p}), b_{s'}(\vec{p'})\} (v_s(p))_\alpha (\bar{v}_{s'}(p'))_\beta e^{i(p\cdot x - p'\cdot y)}]. \qquad (3.225)
\end{aligned}$$

We have omitted the anticommutators that vanish. The remaining two are equal, and given by Equation (3.213) (an anticommutator is invariant under the interchange of the two operators). So, we set $x_0 = y_0$ and integrate over d^3p' to get:

$$\begin{aligned}
\{\psi_\alpha(\vec{x}, t), \bar{\psi}_\beta(\vec{y}, t)\} = \frac{1}{(2\pi)^3} \sum_s \int \frac{d^3p}{2E} &[(u_s(p))_\alpha (\bar{u}_s(p))_\beta e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \\
&+ (v_s(p))_\alpha (\bar{v}_s(p))_\beta e^{i\vec{p}\cdot(\vec{x}-\vec{y})}]. \qquad (3.226)
\end{aligned}$$

The completeness relations for the spinors, Equations (2.160) and (2.161) can be written in component form:

$$\sum_s [u_s(p)\bar{u}_s(p)]_{\alpha\beta} = \gamma^0_{\alpha\beta} E - \vec{\gamma}_{\alpha\beta} \cdot \vec{p} + m\delta_{\alpha\beta}, \qquad (3.227)$$

$$\sum_s [v_s(p)\bar{v}_s(p)]_{\alpha\beta} = \gamma^0_{\alpha\beta} E - \vec{\gamma}_{\alpha\beta} \cdot \vec{p} - m\delta_{\alpha\beta}. \qquad (3.228)$$

Only the $\gamma^0 E$ term survives the integration, leaving us with:

$$\{\psi(\vec{x}, t), \bar{\psi}(\vec{y}, t)\} = \gamma^0 \delta(\vec{x} - \vec{y}). \qquad (3.229)$$

One last formula of interest here is the time ordered product for the Dirac field, that is the electron propagator. First note that because $\psi(x)$ cross couples electron and positron operators, $\langle 0|\, \psi(x)\psi(y)\, |0\rangle = 0$. The product $\psi(x)\psi(y)$ contains the following terms:

$$\psi(x)\psi(y) \sim aa' + ab'^\dagger + b^\dagger a' + b^\dagger b'^\dagger, \qquad (3.230)$$

and the vacuum expectation value of each term vanishes. So, we must look at $\psi(x)\bar{\psi}(y)$. Here we have a non-zero operator product:

$$\psi(x)\bar{\psi}(y) \sim aa'^{\dagger} + ab' + b^{\dagger}a'^{\dagger} + b^{\dagger}b', \tag{3.231}$$

and the first term will give an electron contribution to the vacuum expectation value. Note that the reverse order, $\bar{\psi}(y)\psi(x)$, will give a positron contribution $b'b^{\dagger}$. Because of the column vector–row vector product, it is again useful to work with components:

$$\langle 0|\, \psi_{\alpha}(x)\bar{\psi}_{\beta}(y)\, |0\rangle = \frac{1}{(2\pi)^3}\sum_{s}\int \frac{d^3p}{2E}(u_s(p))_{\alpha}(\bar{u}_s(p))_{\beta}e^{-ip\cdot(x-y)}, \tag{3.232}$$

where we have isolated the non-zero term, and used the anticommutation relation $\{a_s(\vec{p}), a^{\dagger}_{s'}(\vec{p'})\} = (2\pi)^3 2E\delta_{ss'}\delta(\vec{p}-\vec{p'})$. Now use the completeness relation:

$$\sum_{s=1,2}(u_s(p))_{\alpha}(\bar{u}_s(p))_{\beta} = (\gamma^0)_{\alpha,\beta}E - (\vec{\gamma})_{\alpha,\beta}\cdot\vec{p} + m\delta_{\alpha,\beta}, \tag{3.233}$$

to obtain the result:

$$\langle 0|\, \psi_{\alpha}(x)\bar{\psi}_{\beta}(y)\, |0\rangle = \frac{1}{(2\pi)^3}\int \frac{d^3p}{2E}((\gamma^{\mu})_{\alpha,\beta}\, p_{\mu} + m\, \delta_{\alpha,\beta})e^{-ip\cdot(x-y)}. \tag{3.234}$$

What does it mean to reverse the order?

$$\langle 0|\, \bar{\psi}_{\beta}(y)\psi_{\alpha}(x)\, |0\rangle = \frac{1}{(2\pi)^3}\sum_{s}\int \frac{d^3p}{2E}(v_s(p))_{\alpha}(\bar{v}_s(p))_{\beta}e^{ip\cdot(x-y)}, \tag{3.235}$$

and the completeness relation now reads:

$$\sum_{s=1,2}(v_s(p))_{\alpha}(\bar{v}_s(p))_{\beta} = (\gamma^0)_{\alpha,\beta}E - (\vec{\gamma})_{\alpha,\beta}\cdot\vec{p} - m\, \delta_{\alpha,\beta}. \tag{3.236}$$

Inserting the completeness sum into Equation (3.235) gives us

$$\langle 0|\, \bar{\psi}_{\beta}(y)\psi_{\alpha}(x)\, |0\rangle = \frac{1}{(2\pi)^3}\int \frac{d^3p}{2E}((\gamma^{\mu})_{\alpha,\beta}\, p_{\mu} - m\, \delta_{\alpha,\beta})e^{ip\cdot(x-y)}. \tag{3.237}$$

If we reverse the four momentum $p \to -p$, and change the overall sign with $E \to -E$, we turn Equation (3.237) into Equation (3.234).

The contour integration to reduce the four-dimensional form to this expression with the factor $1/2E$ is the same as described in the scalar case. In complex energy space, complete the contour downwards if $x_0 > y_0$, picking up the pole at $E = +E_0$, and multiplying by $-2\pi i$, and reverse the contour for $y_0 > x_0$. The Feynman propagator in configuration space is usually written:

$$\begin{aligned} S_F(x-y) &= \langle 0| T\psi(x)\bar{\psi}(y) |0\rangle \\ &= \frac{1}{(2\pi)^4} \int d^4p e^{-ip\cdot(x-y)} \left\{ \frac{i(\not{p}+m)}{p^2 - m^2 + i\epsilon} \right\}. \end{aligned} \tag{3.238}$$

The propagator in momentum space is in the curly brackets. The interchange of the Fermion field operators requires a minus sign:

$$\begin{aligned} \langle 0| T\psi(x)\bar{\psi}(y) |0\rangle &= \langle 0| \psi(x)\bar{\psi}(y) |0\rangle \, x_0 > y_0 \\ &= -\langle 0| \bar{\psi}(y)\psi(x) |0\rangle \, y_0 > x_0. \end{aligned} \tag{3.239}$$

Another application of time ordering occurs in the solution to the first-order differential equation Equation (3.148) for the unitary time displacement operator in the interaction representation $U_I(t, t_0)$. A formal solution satisfying the boundary condition $U_I(t_0, t_0) = 1$ is the integral equation:

$$U_I(t, t_0) = 1 + \frac{1}{i\hbar} \int_{t_0}^{t} dt' H_I(t') U_I(t', t_0). \tag{3.240}$$

Now assume that H_I is proportional to a dimensionless parameter λ that is small, so that a power series in H_I goes like $1 + \lambda + \lambda^2 + \lambda^3 + \cdots$, and converges rapidly. Think of λ as the elementary charge e, which in natural units is about $e = 0.3$. Then the integral equation can be approximately solved by iteration:

$$\begin{aligned} U_I(t, t_0) = 1 &+ \frac{1}{i\hbar} \int_{t_0}^{t} dt' H_I(t') - \frac{1}{\hbar^2} \int_{t_0}^{t} dt' H_I(t') \int_{t_0}^{t'} dt'' H_I(t'') \\ &- \frac{1}{i\hbar^3} \int_{t_0}^{t} dt' H_I(t') \int_{t_0}^{t'} dt'' H_I(t'') \int_{t_0}^{t''} dt''' H_I(t''') + \cdots, \end{aligned} \tag{3.241}$$

where on the right-hand side $U_I(t_0,t_0)=1$. The nested integrals are time ordered. The upper limit of each integral can be replaced by t provided that the time ordering operator T is placed in front. The factorials take care of the permutations. Thus,

$$\begin{aligned}U_I(t,t_0)=T\Bigg\{1+\frac{1}{i\hbar}\int_{t_0}^{t}dt'H_I(t')-\frac{1}{\hbar^2 2!}\int_{t_0}^{t}dt'H_I(t')\int_{t_0}^{t}dt''H_I(t'')\\ -\frac{1}{i\hbar^3 3!}\int_{t_0}^{t}dt'H_I(t')\int_{t_0}^{t}dt''H_I(t'')\int_{t_0}^{t}dt'''H_I(t''')+\cdots\Bigg\}.\end{aligned}\tag{3.242}$$

This expansion can be written in compact form:

$$U_I(t,t_0)=Te^{\left\{\frac{1}{i\hbar}\int_{t_0}^{t}dt'H_I(t')\right\}}.\tag{3.243}$$

3.13. Problems

One-dimensional harmonic oscillator with Planck's constant restored, so our formulas look like everyone else's.

$$H=\frac{p^2}{2m}+\frac{kx^2}{2}=\frac{\hbar\omega}{2}(P^2+Q^2),$$

$$x_0^2=\frac{\hbar}{m\omega},\ Q=\frac{x}{x_0},\ P=\frac{x_0p}{\hbar},$$

$$a=\frac{Q+iP}{\sqrt{2}},\ a^\dagger=\frac{Q-iP}{\sqrt{2}},$$

$$[x,p]=i\hbar,\ [Q,P]=i,\ \left[a,a^\dagger\right]=1,$$

$$H=\hbar\omega\left(a^\dagger a+\frac{1}{2}\right).$$

Differential equation form

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_n(x)}{dx^2}+\frac{k}{2}x^2\psi_n(x)=E_n\psi_n(x),$$

$$\psi_n(x,t)=e^{-i\omega t/2}\left(\frac{1}{\sqrt{\pi}2^n n!x_0}\right)^{1/2}H_n(x/x_0)e^{-x^2/(2x_0^2)}e^{-in\omega t},$$

The Generating Function for the Hermite Polynomials is

$$e^{\xi^2-(s-\xi)^2} = \sum_{n=0}^{\infty} \frac{H_n(\xi)s^n}{n!}.$$

Taking n derivatives on each side and setting $s = 0$ gives Rodriguez's formula for the polynomials:

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{\partial^n}{\partial \xi^n} e^{-\xi^2}.$$

3.1 Use the commutation rules and the hypothesis $a\,|0\rangle = 0$ to obtain the matrix elements $a\,|n\rangle = \sqrt{n}\,|n-1\rangle$, $a^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle$, and the number operator $N = a^\dagger a$, $N\,|n\rangle = n\,|n\rangle$. Show that $|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}\,|0\rangle$.

3.2 Use $a\,|0\rangle = 0$ in operator form to obtain a first-order differential equation for the ground state wave function $\psi_0(x)$. Solve for $\psi_0(x)$, and normalize it. The Gaussian normalization integral is $\int dx e^{-\frac{x^2}{2\sigma^2}} = \sqrt{2\pi}\sigma$.

3.3 For the harmonic oscillator $\int dx \psi_n^*(x) x \psi_n(x) = \int dx \psi_n^*(x) (d/dx)\psi_n(x) = 0$ for all energy eigenfunctions. The eigenfunctions have definite parity, and the operators have odd parity. Calculate $\int dx \psi_0^*(x) x^2 \psi_0(x)$ and $-\int dx \psi_0^*(x) \frac{d^2}{dx^2} \psi_0(x)$, and show that the uncertainty product $\Delta x \Delta p = \hbar/2$, the minimum uncertainty. Kinetic and potential energy on the average share equally in the energy of the state $E_n = \hbar\omega(n+1/2)$. Use this to obtain a formula for any energy eigenvalue $\langle \Delta x \Delta p \rangle_n = \hbar(n+1/2)$. The general formula with $n = 0$ serves as a check on the wave function calculations. Harmonic oscillator energy eigenfunctions do not oscillate. They are constant in time.

3.4 It is possible to construct an oscillating Gaussian packet that moves back and forth with the classical frequency, and maintains a constant minimum uncertainty. Such a packet contains a distribution of energies, and is called a *coherent state.* Coherent states have a number of unusual properties, and are an important part of quantum optics.

(a) Take a minimum packet, the ground state wave function, and displace it from the origin by a distance $x \to (x - a)$. This is the wave function at $t = 0$:

$$\psi(x, t = 0) = \left(\frac{1}{\sqrt{\pi} x_0}\right)^{1/2} e^{\frac{(x-a)^2}{2x_0^2}} = \sum_{n=0}^{\infty} A_n \psi_n(x), \tag{3.244}$$

The expansion coefficients A_n are constants, the eigenfunctions of the Hamiltonian are $\psi_n(x)$, and the time dependent wave function $\psi(x, t)$ is obtained by multiplying each eigenfunction by $e^{-i\omega(n+1/2)t}$. The coefficients A_n may be determined using the orthonormality of the eigenfunctions:

$$\begin{aligned} A_n &= \int_{-\infty}^{\infty} dx \psi_n^*(x) \psi(x, 0) \\ &= \frac{1}{\sqrt{\pi}} \left(\frac{1}{2^n n!}\right)^{1/2} \int d\xi H_n(\xi) e^{(-\xi^2 + \xi\xi_0 - \xi_0^2/2)}, \end{aligned} \tag{3.245}$$

where the dimensionless variable $\xi = x/x_0$, and the constant $\xi_0 = a/x_0$. The integral looks hopeless, but can be evaluated by using the generating function, and looking at a the following expression:

$$\int d\xi e^{(-s^2 + 2s\xi - \xi^2 + \xi\xi_0 - \xi_0^2/2)} = \sum_{n=0}^{\infty} \int d\xi \frac{H_n(\xi) s^n}{n!} e^{(-\xi^2 + \xi\xi_0 - \xi_0^2/2)}. \tag{3.246}$$

The author is indebted to Schiff's '*Quantum Mechanics*' for this analysis. The integral on the left can be evaluated by completing the square in the exponential, and using the Gaussian integral formula. Show that the result is

$$\int d\xi e^{(-s^2 + 2s\xi - \xi^2 + \xi\xi_0 - \xi_0^2/2)} = \sqrt{\pi} e^{-\xi_0^2/4} e^{s\xi_0}. \tag{3.247}$$

Then expand $e^{s\xi_0}$ in a Taylor series and match orders of s on both sides of the equation to obtain:

$$A_n = \left(\frac{1}{2^n n!}\right)^{1/2} e^{-\xi_0^2/4} (\xi_0)^n. \tag{3.248}$$

This gets substituted into Equation (3.244). Add in the energy-time phases to get $\psi(x,t)$. Now use the generating function in reverse to eliminate $H_n(\xi)$, by identifying $s = \frac{\xi_0 e^{-i\omega t}}{2}$. Show that the oscillating wave function is now a pure exponential:

$$\psi(x,t) = \left(\frac{x_0}{\sqrt{\pi}}\right)^{1/2} \exp\left(-\xi^2/2 - \frac{\xi_0^2}{4}(1+\cos(2\omega t)) + \xi\xi_0\cos(\omega t) \right.$$
$$\left. + i\left(\frac{\xi_0^2 \sin(2\omega t)}{4} - \xi_0\xi\sin(\omega t) - \omega t/2\right)\right). \quad (3.249)$$

Show that $\psi(x, t = 0)$ is an eigenfunction of the lowering (destruction) operator $a = (Q + iP)/(\sqrt{2})$ with eigenvalue $\xi_0/\sqrt{2}$. Then obtain the magnitude squared of the oscillating wave function:

$$|\psi(x,t)|^2 = \frac{1}{x_0\sqrt{\pi}} e^{-\frac{(x-a\cos(\omega t))^2}{x_0^2}}. \quad (3.250)$$

This is a Gaussian shaped minimum uncertainty wave packet that oscillates back and forth about the origin with amplitude a. It was obtained simply by displacing the ground state wave function a distance a from the origin, and letting it loose.

(b) Coherent states in ket language. Define the eigenket $|\lambda\rangle$:

$$|\lambda\rangle = e^{-|\lambda|^2/2} e^{\lambda a^\dagger} |0\rangle, \quad (3.251)$$

where the eigenvalue λ is in general a complex number, since the operator a is not Hermitian. Show that $|\lambda\rangle$ is an eigenket of a:

$$a|\lambda\rangle = \lambda|\lambda\rangle. \quad (3.252)$$

Use the power series expansion of the exponential to show that:

$$|\lambda\rangle = e^{-|\lambda|^2/2} \sum_{n=0}^{\infty} \frac{\lambda^n |n\rangle}{\sqrt{n!}}. \quad (3.253)$$

By comparing this power series with Equation (3.244), we have $\xi_0 = \lambda\sqrt{2}$. Defining $\langle\lambda|$ from the power series for the ket:

$$\langle\lambda| = e^{-|\lambda|^2/2} \sum_{n'=0}^{\infty} \frac{(\lambda^*)^{n'} \langle n'|}{\sqrt{n'!}}. \quad (3.254)$$

Use these forms to show that $|\lambda\rangle$ is normalized:

$$\langle\lambda|\lambda\rangle = 1. \tag{3.255}$$

The normalization can also be checked from the conjugate form of Equation (3.251), namely,

$$\langle\lambda| = e^{-|\lambda|^2/2}\,\langle 0|\, e^{\lambda^* a}. \tag{3.256}$$

The proof of this assertion is deferred until Wick's theorem has been introduced. See Problem 3.10.

Derive the average energy level quantum number $\bar{n} = |\lambda|^2$ using Equations (3.253) and (3.254), and the operator $N = a^\dagger a$. Show that the eigenvalues are Poisson distributed:

$$Pr\{n\} = \frac{\bar{n}^n e^{-\bar{n}}}{n!}. \tag{3.257}$$

3.5 If someone told you that the complete angular momentum algebra can be derived from two independent, non-interacting one-dimensional harmonic oscillators, you might be surprised, but you might also guess that it is not too difficult to prove. If you know that something can be done, it is often easy to do it. Schwinger was the first to show this. Take the angular momentum matrix elements in the following forms:

$$\begin{aligned} J_+ &= J_x + iJ_y; J_+\,|j,m\rangle \\ &= \sqrt{(j-m)(j+m+1)}\,|j,m+1\rangle, \end{aligned} \tag{3.258}$$

$$\begin{aligned} J_- &= J_+^\dagger = J_x - iJ_y; J_-\,|j,m\rangle \\ &= \sqrt{(j+m)(j-m+1)}\,|j,m-1\rangle, \end{aligned} \tag{3.259}$$

$$J_z\,|j,m\rangle = m\,|j,m\rangle, \tag{3.260}$$

$$J^2 = (J_+J_- + J_-J_+)/2 + J_z^2; J^2\,|j,m\rangle = j(j+1)\,|j,m\rangle, \tag{3.261}$$

$$[J_+, J_-] = 2J_z, \tag{3.262}$$

where $-j \le m \le +j, 2j+1$ values, j integer or half integer, including 0.

Call the two independent harmonic oscillators a and b, each perfectly ordinary, so $\left[a, a^\dagger\right] = 1$, and $\left[b, b^\dagger\right] = 1$, and everything else commutes, in particular $\left[a, b^\dagger\right] = [a, b] = 0$. Quantum numbers are n_a and n_b, and $a\,|n_a\rangle = \sqrt{n_a}\,|n_a - 1\rangle$, etc. Try $J_+ = a^\dagger b$. Show that the matrix elements for J_+ and J_- are correct if $n_a = j + m$ and $n_b = j - m$. Calculate the matrix elements of J_z and J^2. With $j = (n_a + n_b)/2$ and $m = (n_a - n_b)/2$, show that all representations and eigenvalues for J^2 and J_z are accounted for. Make a table of n_a, n_b, j, m for $j = 0, 1/2, 1, 3/2$. The matrices in the n_a, n_b representation are awkward. To begin with, the labels are upside down, since $n = 0$ is at the top of the infinite matrix, and the raising operator goes down one notch. Nevertheless, one can use the matrices. Write down $2J_z = a^\dagger a - b^\dagger b$, and find the eigenvectors and eigenvalues, for spin 1/2. *Hint*:

$$\left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right] \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{3.263}$$

$a^\dagger a$ is the first 2×2 matrix, and $n_a = 1$ is the ket on the left; $n_b = 0$ is on the right. The a matrices do not see the b matrices or the b kets.

3.6 Using the definition of the field amplitude $\Phi(x)$ in Equation (3.43), show that the commutation rules Equations (3.45)–(3.47) give the desired result $[\Phi(x), \dot{\Phi}(y)] = i\delta(x - y)$.

3.7 Obtain the formula for the Hamiltonian Equation (3.54).

3.8 Charged field. Part 1, Klein–Gordon equation. There are four classical fields in this problem, all with the same mass, and all satisfying the Klein–Gordon equation. Show that the Lagrange density for two non-interacting real fields $\phi_1(x)$ and $\phi_2(x)$:

$$\mathcal{L} = (\partial_\mu \phi_1 \partial^\mu \phi_1 - m^2 \phi_1^2)/2 + (\partial_\mu \phi_2 \partial^\mu \phi_2 - m^2 \phi_2^2)/2, \tag{3.264}$$

becomes one equation in ϕ and ϕ^*:

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi, \tag{3.265}$$

where $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$ and $\phi^* = (\phi_1 - i\phi_2)/\sqrt{2}$. Treat ϕ and ϕ^* as independent variables in Lagrange's equations. Show that the four-dimensional current j^μ is conserved:

$$\partial_\mu(\phi^* \partial^\mu \phi - (\partial^\mu \phi^*)\phi) = 0. \tag{3.266}$$

Define the charge density ρ as the elementary charge times the time component of j^μ:

$$\rho = -iej^0(x) = -ie(\phi^* \partial\phi/\partial t - (\partial\phi^* \partial t)\phi). \tag{3.267}$$

As mentioned in the text, the quantized field (no longer Hermitian) has to couple operators for Φ and $\Phi^\dagger$, thus,

$$\Phi(x) = \frac{1}{(2\pi)^3} \int \frac{d^3k}{2\omega} (a(\vec{k})e^{-ik\cdot x} + b^\dagger(\vec{k})e^{ik\cdot x}). \tag{3.268}$$

The a's and b's satisfy the usual commutation relations Equations (3.45)–(3.47), while a and $a^\dagger$ commute with b and $b^\dagger$. Show that after normal ordering the total charge Q is given by:

$$Q = \int d^3x\rho(x) = \frac{e}{(2\pi)^3} \int \frac{d^3k}{2\omega} (a^\dagger(\vec{k})a(\vec{k}) - b^\dagger(\vec{k})b(\vec{k})). \tag{3.269}$$

Normal ordering reversed the order of the b term, but the minus sign came from the definition of ρ. The time dependent terms in the expansion also vanish because ρ is the difference between two terms.

3.9 Charged field Part 2, Dirac equation. Repeat the calculation of the total charge for the Dirac operator $j^\mu = e\bar{\Psi}\gamma^\mu\Psi$. The result is the same, apart from a sum over spins. Note the differences in the calculation. $\Psi^\dagger\Psi$ has no time derivatives, and is not the difference of two quantities. The cancellation is taken care of by the orthogonality of the spinors, and the sign change on the $b^\dagger b$ term is achieved because b and $b^\dagger$ anticommute.

3.10 Wick's theorem and the harmonic oscillator. We return to the problem of the normalization of coherent states of the harmonic oscillator:

$$|\lambda\rangle = e^{-|\lambda|^2/2} e^{\lambda a^\dagger} |0\rangle , \tag{3.270}$$

and

$$\langle\lambda| = e^{-|\lambda|^2/2} \langle 0| e^{\lambda^* a}. \tag{3.271}$$

Multiplying them together gives:

$$\langle\lambda|\lambda\rangle = e^{-\lambda^2} \langle 0| e^{\lambda a} e^{\lambda a^\dagger} |0\rangle , \tag{3.272}$$

where we have taken λ to be real, which it is in Schrödinger's solution. It is the vacuum expectation value of $e^{\lambda a} e^{\lambda a^\dagger}$ that must equal e^{λ^2} for the normalization. This is true, but proving it is an interesting exercise in the use of Wick's theorem for the reduction of vacuum expectation values. First, because $\left[a, a^\dagger\right] = 1$,

$$e^{\lambda a} e^{\lambda a^\dagger} = e^{\lambda(a+a^\dagger)+\lambda^2\left[a,a^\dagger\right]/2}, \tag{3.273}$$

and half of our problem is taken care of, so we are left with the task of showing that:

$$e^{\lambda^2/2} = \langle 0| e^{\lambda(a+a^\dagger)} |0\rangle . \tag{3.274}$$

Expand the exponential into products of the operators:

$$\begin{aligned} e^{\lambda(a+a^\dagger)} &= 1 + \lambda(a+a^\dagger) + \lambda^2(a+a^\dagger)^2/2! + \lambda^3(a+a^\dagger)^3/3! \\ &\quad + \lambda^4(a+a^\dagger)^4/4! + \lambda^5(a+a^\dagger)^5/5! \\ &\quad + \lambda^6(a+a^\dagger)^6/6! + \cdots . \end{aligned} \tag{3.275}$$

The vacuum expectation values of all odd powers of the operators vanish, so we are left with $0, 2, 4, 6, \ldots$ that match the powers of λ^2 in $e^{\lambda^2/2}$. Make a table matching coefficients of λ^2 in Equation (3.275) through sixth order. To have a non-vanishing vacuum expectation value, there must be an equal number of a's and $a^\dagger$'s; an a must be on the left, and an $a^\dagger$ on the right. So, $(a + a^\dagger)^4$ has two non-vanishing arrangements, namely, $aa^\dagger aa^\dagger$, and $aaa^\dagger a^\dagger$, and there are three individual pairings:

$$\langle 0|\, aa^\dagger aa^\dagger \,|0\rangle = \contraction{}{a}{}{a^\dagger}\contraction{aa^\dagger}{a}{}{a^\dagger}aa^\dagger aa^\dagger = 1, \tag{3.276}$$

$$\langle 0|\, aaa^\dagger a^\dagger \,|0\rangle = \contraction{}{a}{a}{a^\dagger}\contraction[2ex]{a}{a}{a^\dagger}{a^\dagger}aaa^\dagger a^\dagger = 1, \tag{3.277}$$

$$\langle 0|\, aaa^\dagger a^\dagger \,|0\rangle = \contraction{}{a}{aa^\dagger}{a^\dagger}\contraction[2ex]{a}{a}{}{a^\dagger}aaa^\dagger a^\dagger = 1. \tag{3.278}$$

All contractions are either one or zero. For the fourth-order term they add up to 3. The sixth-order term is not as daunting as it may seem. Gather your courage. Multiplying everything out gives 64 terms, but all but six have zero vacuum expectation value. The six terms are of the form $\langle 0|\, axxyya^\dagger \,|0\rangle$ with two a's and two $a^\dagger$'s in the middle, because $4!/(2!)^2 = 6$ (interchanging the two a's or the two $a^\dagger$'s doesn't change anything). One of these terms has no contribution, but the other five do. Work it out. Show that $\langle 0|\, axxyya^\dagger \,|0\rangle = 15$, and that is exactly what is needed to match the two power series. Checking the formula to sixth order will suffice to prove the identity.

3.11 Show that adding the four divergence $\partial_\mu A_0 F^{0\mu}$ to the Hamiltonian density for the electromagnetic field Equation (3.107), gives the desired form $\mathcal{H} = (\vec{E}^2 + \vec{B}^2)/2$. Remember that for a free field $\vec{\nabla} \cdot \vec{E} = 0$.

3.12 Gupta and Bleuler's argument, adopted from Otto Nachtmann's text. We learned that the gauge condition $k^\mu a_\mu = 0$ did not make sense as an operator equation, but it is possible to require:

$$k^\mu a_\mu \,|\text{physical photon state}\rangle = 0. \tag{3.279}$$

Choose the z axis for the photon direction $\vec{k} = \omega \hat{z}$, and the allowed polarization directions are ϵ_x and ϵ_y. Define linear combinations of $a_0(\vec{k})$ and $a_3(\vec{k})$ thus,

$$\alpha_0(\vec{k}) = (a_0(\vec{k}) - a_3(\vec{k}))/\sqrt{2}, \tag{3.280}$$

$$\alpha_3(\vec{k}) = (a_0(\vec{k}) + a_3(\vec{k}))/\sqrt{2}, \tag{3.281}$$

$$\alpha_1(\vec{k}) = a_1(\vec{k}), \tag{3.282}$$

$$\alpha_2(\vec{k}) = a_2(\vec{k}). \tag{3.283}$$

α_0 is chosen so that the gauge condition is:

$$\alpha_0(\vec{k})\,|\text{physical photon state}\rangle = 0. \tag{3.284}$$

$\alpha_3^\dagger(\vec{k})$ creates a rogue state that does not satisfy $\epsilon_\mu k^\mu = 0$. Using the commutation rules for $a_\mu(\vec{k})$ and $a_\mu^\dagger(\vec{k})$ Equations (3.113) and (3.114), work out the commutation rules for the α's and $\alpha^\dagger$'s. Show that a multiphoton state created by a string of $\alpha^\dagger$'s satisfies Equation (3.284) ($\alpha_0\,|0\rangle = 0$), provided that there are no states created by an $\alpha_3^\dagger$.

3.13 Beginning with $\vec{A}$ in Equation (3.128), go through the necessary steps to obtain Equation (3.134), the electromagnetic Hamiltonian:

$$H = \frac{1}{(2\pi)^3}\int \frac{d^3k}{2}\sum_s a^{\dagger(s)}(\vec{k})a^{(s)}(\vec{k}). \tag{3.285}$$

3.14 Derive Equation (3.148).

3.15 Check the integration to obtain

$$\langle f|\,H_I\,|i\rangle = \frac{ZZ'e^2}{V}\frac{4\pi}{(K^2+\alpha^2)}. \tag{3.286}$$

3.16 Derive the Green's function $G_k(\vec{x},\vec{x'})$ Equation (3.174) from Equation (3.172).

3.17 Obtain Equation (3.183):

$$[\Phi(x),\Phi(y)] = \frac{1}{(2\pi)^3}\int \frac{d^3k}{2\omega}\{e^{-ik\cdot(x-y)} - e^{ik\cdot(x-y)}\}. \tag{3.287}$$

3.18 Verify Equation (3.197) for the photon propagator.

3.19 Show that Equation (3.204) is gauge invariant with $\psi \to \psi e^{ie\chi}$.

3.20 Obtain Equation (3.223) for the Dirac Hamiltonian.

References

[1] O. Nachtmann, *Elementary Particle Physics* (Springer-Verlag, Berlin, 1990).
[2] M. Peskin and D. Schroeder, *Introduction to Quantum Field Theory* (Addison-Wesley, New York, 1995).
[3] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, New York, 1949), p. 58.

This page intentionally left blank

Chapter 4

Quantum Electrodynamics

4.1. *S* and *T* Matrices

The calculation of transition rates between initial and final states caused by an interaction Hamiltonian $H_I(t)$ is based on the expansion of $U_I(t, t_0)$ in Equation (3.239), with the limits extended to $\pm\infty$. Thus,

$$\begin{aligned} &U_I(\infty, -\infty) \\ &\quad = T\left\{1 + \frac{1}{i\hbar}\int_{-\infty}^{\infty} dt' H_I(t') - \frac{1}{\hbar^2 2!}\int_{-\infty}^{\infty} dt' H_I(t')\int_{-\infty}^{\infty} dt'' H_I(t'')\right. \\ &\qquad \left. - \frac{1}{i\hbar^3 3!}\int_{-\infty}^{\infty} dt' H_I(t')\int_{-\infty}^{\infty} dt'' H_I(t'')\int_{-\infty}^{\infty} dt''' H_I(t''') + \cdots\right\}. \end{aligned} \tag{4.1}$$

$U_I(\infty, -\infty)$ is called the S matrix. The probability amplitude for the Hamiltonian H_I to cause a transition from initial state $|i\rangle$ to final state $|f\rangle$ is the matrix element:

$$\text{Amplitude } i \to f = \langle f|\, S\, |i\rangle \tag{4.2}$$

The one in the expansion Equation (4.1) means nothing happens, and becomes a Kronecker delta $\delta_{i,f}$. The T matrix is defined as that part of the S matrix for which the final state differs from the initial state, and includes a four-dimensional delta function to conserve momentum and energy:

$$\langle f|\, S\, |i\rangle = \delta_{i,f} + i(2\pi)^4\delta^4(p_f - p_i)\, \langle f|\, T\, |i\rangle. \tag{4.3}$$

The calculation of T matrix elements is the task of covariant perturbation theory. We need a technique for handling the continuum normalization of the state vectors. From Chapter 3 we may write for an electron wave function:

$$\langle 0|\, a_s(\vec{p})a^\dagger_{s'}(\vec{p'})\, |0\rangle = (2\pi)^3 2E\delta_{s,s'}\delta(\vec{p} - \vec{p'}). \tag{4.4}$$

Now for a bit of magic, attributed to Fermi (probably thoroughly justified). We write the δ function as an integral:

$$\delta(\vec{p} - \vec{p'}) = \frac{1}{(2\pi)^3}\int d^3x e^{i(\vec{p}-\vec{p'})\cdot\vec{x}}, \tag{4.5}$$

which for $\vec{p} = \vec{p'}$ becomes an integral over the space volume:

$$\langle 0|\, a_s(\vec{p})a^\dagger_s(\vec{p})\, |0\rangle = 2EV, \tag{4.6}$$

where we have dropped the primes — the two states are the same. So, we have acquired a normalization volume, just like the box normalized wave functions in Chapter 3. Without this trick the expression in Equation (4.6) would diverge. The final calculations must be of course independent of the volume, as we shall see.

Perhaps a word about dimensions is appropriate. In our units $\hbar = 1$ and $c = 1$, so length and time have the same dimensions, and energy is 1/time, so we may work with length and 1/length. The operator S is dimensionless, but the dimensions of the matrix element $\langle f|\, S\, |i\rangle$ depend on the number of particles. For $2 \to 2$ $\langle f|\, S\, |i\rangle \sim l^4$. Since the four-dimensional δ function in the definition of the T matrix $\delta(p_f - p_i) \sim l^4$ also, the T matrix element for this case is dimensionless.

4.2. Transition Rates and Cross-Sections

The transition rate $w_{i\to f}$ must contain the square of the T matrix element, the final state phase space, a δ function to conserve momentum and energy, and numerical factors for state vector normalization.

When we square the S matrix element, we are confronted by the square of a δ function, that we handle again using Fermi's trick:

$$|(2\pi)^4\delta^4(p_f - p_i)\,\langle f|\,T\,|i\rangle\,|^2 = (2\pi)^4\delta^4(p_f - p_i)|\,\langle f|\,T\,|i\rangle\,|^2(Vt). \tag{4.7}$$

The last factor (Vt) comes from the integral over space–time after setting the phase in the exponential to zero. One factor of $(2\pi)^4$ has been absorbed by the δ function. The time comes in handy to convert the probability into a rate. Putting it all together, and dividing by t gives the transition rate formula for two particles in the initial state going to two particles in the final state. The final state particles can be the same as the initial state, like $e^- + e^+ \to e^- + e^+$, or different, like $e^- + e^+ \to \gamma + \gamma$:

$$\begin{aligned} dw_{i\to f} &= (2\pi)^4\delta^4(p_1 + p_2 - p_3 - p_4)\frac{d^3p_3}{(2\pi)^3 2E_3}\frac{d^3p_4}{(2\pi)^3 2E_4} \\ &\qquad \times |\,\langle f|\,T\,|i\rangle\,|^2\frac{1}{2E_1V}\frac{1}{2E_2V}V. \end{aligned} \tag{4.8}$$

Let us check dimensions again here. As explained above, the T matrix element is dimensionless, and the dimensions of the four-dimensional δ function and the phase space factors cancel each other, leaving $dw_{i\to f} \sim 1/(E^2V) = 1/l = 1/\text{time}$, and that is correct!

The cross-section is the transition rate divided by the flux, and the flux will cancel the remaining normalization volume. In Chapter 3, Rutherford's formula was derived using the α particle flux $= \rho v = v/V$. The relativistic expression is not that different. We imagine the initial state particles to be coming towards each other, like in a colliding beam storage ring. The relative velocity is then:

$$v_{\text{rel}} = \frac{|\vec{p_1}|}{E_1} + \frac{|\vec{p_2}|}{E_2}. \tag{4.9}$$

Note that this expression can exceed the speed of light. In fact, two highly relativistic beams, one of electrons, the other of positrons, pass through each other with speed $2c$. Dividing Equation (4.8) by

v_{rel}/V gives the following expression for the differential cross-section:

$$d\sigma_{i\to f} = (2\pi)^4\delta^4(p_1+p_2-p_3-p_4)\frac{d^3p_3}{(2\pi)^3 2E_3}\frac{d^3p_4}{(2\pi)^3 2E_4}$$
$$\times\,|\,\langle f|\,T\,|i\rangle\,|^2\frac{1}{4(E_2|\vec{p_1}|+E_1|\vec{p_2}|)}. \tag{4.10}$$

The dimensions are given by the flux term $\sim 1/(E^2)$. Here the conversion factor $1(\text{GeV})^{-2} = 0.389$ millibarns comes in handy. There are several equivalent expressions for the flux. It is not relativistically invariant in general — only for Lorentz transformations along the beam direction. It is an interesting exercise in kinematics to show that:

$$4(E_2|\vec{p_1}|+E_1|\vec{p_2}|) = 4\sqrt{(p_1\cdot p_2)^2 - m_1^2m_2^2}, \tag{4.11}$$

and

$$4\sqrt{(p_1\cdot p_2)^2 - m_1^2m_2^2} = 2\sqrt{\lambda(s,m_1^2,m_2^2)}, \tag{4.12}$$

where $\lambda(x,y,z) = x^2+y^2+z^2-2xy-2xz-2yz$ was defined in Chapter 2 after Equation (2.64) for the momentum squared of either particle in the center of mass of a two body decay. Equation (4.11) is not true in general — only for the special case that $\vec{p_1}$ and $\vec{p_2}$ are oppositely directed, so that $-\vec{p_1}\cdot\vec{p_2} = +|\vec{p_1}||\vec{p_2}|$. Equation (4.12) is not restricted. If there are more than two particles in the final state, multiply by the factor $\frac{d^3p}{(2\pi)^3 2E}$ for each one, and add its four momentum into the δ function. The normalization volume that appeared in the initial state has already been included in the final state phase space factor. The cross-section is left as a differential. The four-dimensional δ function has to be integrated out, taking care of four differentials in the phase space. The remaining phase space differentials can be converted into a solid angle $d\Omega$ to give $d\sigma/d\Omega$. The T matrix element is in general a function of the energy and the scattering angle. If the particles have spin the T matrix element will depend on the spins. For unpolarized particles in the initial state producing particles in the final state for which the spin is not measured, the correct procedure is to average over the initial spins, and sum over the final spins.

Equation (4.10) is an area perpendicular to the beam direction, and as such is invariant to Lorentz transformations along the axis of the two beams. Another formula of interest that can be derived from Equation (4.8) is the decay rate for an unstable particle. In this case there is a natural frame of reference, namely the particle rest frame. The decay rate into a specific final state is written $d\Gamma_{i\to f}$, for the decay $m \to m_1 + m_2 + m_3 + \cdots$, where we allow for an arbitrary number of final state particles. As we learned in Chapter 2, three particle final states are very common.

$$d\Gamma_{i\to f} = \frac{(2\pi)^4}{2m}\delta(m - E_1 - E_2 - E_3 - \cdots)\delta(\vec{p_1} + \vec{p_2} + \vec{p_3} + \cdots) \\ \times |\langle f|\, T\, |i\rangle|^2 \prod_{i=1}^{n} \frac{d^3p_i}{(2\pi)^3 2E_i}. \tag{4.13}$$

Integration over all of the final state variables gives the decay rate into this particular final state, $K^+ \to \pi^+\pi^+\pi^-$ for instance. The reciprocal of the lifetime of the unstable particle m is the sum of decay rates to all possible final states:

$$\Gamma_i = \sum_f \Gamma_{i\to f} = 1/\tau. \tag{4.14}$$

Γ is usually measured in energy units, corresponding to a line width on a plot of the parent mass.

4.3. $e^- + e^+ \to \mu^- + \mu^+$

We are ready to solve a real problem! The choice of the $\mu^-\mu^+$ final state rather than e^-e^+ avoids having to work with two Feynman diagrams at the price of introducing a new particle, the muon. Muons are members of the lepton family that were introduced in Chapter 1. Table 4.1 shows the three known lepton families, with separately conserved lepton numbers assigned to negatively charged leptons and their associated neutrinos. An identical table could be written for the antileptons, with opposite charge (e^+, μ^+, τ^+), same mass, spin,

Table 4.1: Leptons.

Name	Symbol	Mass (MeV)	Spin	Interactions	Lifetime (s)	Decaymodes	Lepton number
Electron	e^-	0.511	1/2	em, weak	Stable		$l_e = +1$
E neutrino	ν_e	0	1/2	weak	Stable		$l_e = +1$
Muon	μ^-	105.6	1/2	em, weak	2.2×10^{-6}	$e\bar{\nu}_e\nu_\mu$	$l_\mu = +1$
Mu Neutrino	ν_μ	0	1/2	weak	Stable		$l_\mu = +1$
Tau	τ^-	1777.0	1/2	em, weak	2.9×10^{-13}	$\mu\bar{\nu}_\mu\nu_\tau, \pi^-\pi^0\nu_\tau$	$l_\tau = +1$
Tau Neutrino	ν_τ	0	1/2	weak	Stable		$l_\tau = +1$

and lifetime, but lepton number -1. Neutrino mixing violates the conservation of lepton numbers, and also means that neutrino masses cannot all be exactly zero. These small but important effects are ignored in the table.

Figure 1.8 shows the two diagrams that are active for $e^-e^+ \to e^-e^+$, the first is called 't channel' for the photon exchange between the electron and positron, and the second is called 's channel', where the initial electron and positron annihilate to produce an intermediate photon that creates the outgoing electron–positron pair. Four momentum is conserved at each vertex, so the t channel photon has $q^2 = (p_1 - p_3)^2 < 0$, and the s channel photon has $q^2 = (p_1 + p_2)^2 > 0$, meaning that neither photon is real. In fact in the center of mass the s channel photon has only energy and no momentum. Note that the positron lines are drawn backwards — an incoming positron looks like an outgoing electron. For the $\mu^-\mu^+$ final state only the s channel diagram is viable. An electron absorbing a photon and turning into a muon at a single vertex is forbidden. (This would allow the decay $\mu^- \to e^-\gamma$ which violates lepton number conservation and has an experimental upper limit $\Gamma(\mu \to e\gamma)/\text{all} < 5.7 \times 10^{-13}$ [1]). The charged leptons and their neutrinos are assumed to obey the Dirac equation with the appropriate mass term. This is consistent with experiment. There is no evidence for substructure in the leptons, in contrast with the hadrons that are composed of quarks and gluons.

The electron field operator is given by Equation (3.205), reproduced here for convenience:

$$\psi(x) = \frac{1}{(2\pi)^3} \int \frac{d^3p}{2E} \sum_{s=1,2} \left\{ a_s(\vec{p}) u_s(p) e^{-ip\cdot x} + b_s^\dagger(\vec{p}) v_s(p) e^{ip\cdot x} \right\}, \tag{4.15}$$

and

$$\bar{\psi}(x) = \frac{1}{(2\pi)^3} \int \frac{d^3p'}{2E'} \sum_{s'=1,2} \left\{ a_{s'}^\dagger(\vec{p}\,') \bar{u}_{s'}(p') e^{ip'\cdot x} + b_{s'}(\vec{p}\,') \bar{v}_{s'}(p') e^{-ip'\cdot x} \right\}. \tag{4.16}$$

The muon field operator has the same form, but they are different particles, so we must keep them separate. There are already lots of μ's in our equations, so we shall call the muon operator $\psi_m(x')$, and write the operators as $c_r(\vec{q})$ and $d_r(\vec{q})$. The electron and muon operators anticommute to zero — they are in different spaces.

$$\psi_m(x') = \frac{1}{(2\pi)^3} \int \frac{d^3q}{2E_q} \sum_{r=1,2} \left\{ c_r(\vec{q}) u_r^m(q) e^{-iq\cdot x'} + d_r^\dagger(\vec{q}) v_r^m(q) e^{iq\cdot x'} \right\}, \tag{4.17}$$

and

$$\begin{aligned} \bar{\psi}_m(x') = \frac{1}{(2\pi)^3} \int \frac{d^3q'}{2E_q'} \sum_{r'=1,2} & \left\{ c_{r'}^\dagger(\vec{q}\,') \bar{u}_{r'}^m(q') e^{iq'\cdot x'} \right. \\ & \left. + d_{r'}(\vec{q}\,') \bar{v}_{r'}^m(q') e^{-iq'\cdot x'} \right\}. \end{aligned} \tag{4.18}$$

$\bar{\psi} = \psi^\dagger \gamma^0$ in each case, and the notation p, x in ψ, p', x in $\bar{\psi}$, q, x' in ψ_m, and q', x' in $\bar{\psi}_m$ will be preserved throughout the calculation. The interaction Hamiltonian occurs twice in our diagram, once at each vertex. The natural time ordering would be first the e^-e^+ annihilation to produce the intermediate photon that then creates the final state muon pair. However, quantum mechanics allows the opposite order, in which the muon pair and photon are created from the vacuum, and the e^-e^+ pair rushes in to annihilate and supply the necessary energy. The two time ordered processes are in fact the same, and that cancels the 2! in the denominator of the S matrix. We write the S matrix Equation (4.1) as follows:

$$S = U_I(\infty, -\infty) = T\left\{-\frac{1}{2!}\int_{-\infty}^{\infty} dt H_I(t)\int_{-\infty}^{\infty} dt' H_I(t')\right\}, \quad (4.19)$$

where

$$H_I(t) = -e\int d^3x : \bar{\psi}(\vec{x}, t)\gamma^\mu\psi(\vec{x}, t) : A_\mu(\vec{x}, t), \quad (4.20)$$

and

$$H_I(t') = -e\int d^3x' : \bar{\psi_m}(\vec{x'}, t')\gamma^\nu\psi_m(\vec{x'}, t') : A_\nu(\vec{x'}, t'). \quad (4.21)$$

The initial state is an electron with spin s_1 and momentum $\vec{p_1}$ that annihilates with a positron of spin s_2 and momentum $\vec{p_2}$, and the final state is a μ^- with spin s_3 and momentum $\vec{p_3}$, and a μ^+ with spin s_4 and momentum $\vec{p_4}$. The square of the total energy is $s = (p_1 + p_2)^2 > 4m_\mu^2$. The scattering matrix element is

$$\langle f|\, S\, |i\rangle = \langle 0|\, c_{s_3}(\vec{p_3})d_{s_4}(\vec{p_4})S a_{s_1}^\dagger(\vec{p_1})b_{s_2}^\dagger(\vec{p_2})\,|0\rangle. \quad (4.22)$$

There are no real photons in either the initial or final states, so no photon field operators appear in the vacuum expectation value. The whole expression is rather lengthy. There are two integrals over four-dimensional space time: $\int d^4x$ and $\int d^4x'$, that come from the volume and time integrals of the Hamiltonian density. There are four integrals over space momentum, $\int d^3p/(2E)$ from $\psi(x)$, $\int d^3p'/(2E')$ from $\bar{\psi}(x)$, $\int d^3q/(2E_q)$ from $\psi_m(x')$, and finally $\int d^3q'/(2E_q')$ from $\bar{\psi}_m(x')$. A fifth momentum integral will be supplied by the photon propagator. The vacuum expectation value in Equation (4.22) involves the field operators in $\psi(x)$, etc. The usual rules apply, so that only the product $a_s(\vec{p})b_{s'}(\vec{p}')$ contributes for the electrons, and only $c_{r'}^\dagger(\vec{q}')d_r^\dagger(\vec{q})$ for the muons. The vacuum expectation values couple the continuum variables to those of the initial and final states.

The relevant terms in the current densities are:

$$\bar{\psi}(x)\gamma^{\mu}\psi(x) \sim a_s(\vec{p})b_{s'}(\vec{p}\,')\bar{v}_{s'}(p')\gamma^{\mu}u_s(p)e^{-i(p+p')\cdot x}, \tag{4.23}$$

and

$$\bar{\psi}_m(x')\gamma^{\nu}\psi_m(x') \sim c^{\dagger}_{r'}(\vec{q}\,')d^{\dagger}_r(\vec{q})\bar{u}^m_{r'}(q')\gamma^{\nu}v^m_r(q)e^{i(q+q')\cdot x'}. \tag{4.24}$$

The terms $ab^{\dagger}$ and $cd^{\dagger}$ that have to be reversed (with a minus sign) for normal ordering play no role in the vacuum expectation value. The two currents are independent and can be evaluated separately.

Using the anticommutation rules, the vacuum expectation value of the S matrix element is

$$\begin{aligned}\langle f|\,S\,|i\rangle = \frac{-e^2}{(2\pi)^{12}}\int d^4x\int d^4x'\int d^3p/(2E)\int d^3p'/(2E')\sum_{s,s'}\\ \times 2E(2\pi)^3\delta_{s,s_1}\delta(\vec{p}-\vec{p}_1)2E'(2\pi)^3\delta_{s',s_2}\delta(\vec{p}\,'-\vec{p}_2)\bar{v}_{s'}(p')\\ \times\gamma^{\mu}u_s(p)e^{-i(p+p')\cdot x}\int d^3q/(2E_q)\int d^3q'/(2E'_q)\sum_{r,r'}\\ \times 2E_q(2\pi)^3\delta_{r,s_3}\delta(\vec{q}-\vec{p}_3)2E'_q(2\pi)^3\delta_{r',s_4}\delta(\vec{q}\,'-\vec{p}_4)\bar{u}^m_{r'}(q')\\ \times\gamma^{\nu}v^m_r(q)e^{i(q+q')\cdot x'}\,\langle 0|\,T(A_{\mu}(x)A_{\nu}(x'))\,|0\rangle.\end{aligned} \tag{4.25}$$

The two vector potentials $\langle 0|\,T(A_{\mu}(x)A_{\nu}(x'))\,|0\rangle$ are so far just along for the ride. The momentum integrals plus the delta functions, and the spin sums plus the Kronecker δ's replace: $\vec{p}\to\vec{p}_1$, $\vec{p}\,'\to\vec{p}_2$, $\vec{q}\,'\to\vec{p}_3$, $\vec{q}\to\vec{p}_4$, $s\to s_1$, $s'\to s_2$, $r'\to s_3$, and $r\to s_4$. Doing these sums and integrals, and substituting Equation (3.197) for the photon propagator gives:

$$\begin{aligned}\langle f|\,S\,|i\rangle = -e^2\int d^4x\int d^4x'\bar{v}_{s_2}(p_2)\gamma^{\mu}u_{s_1}(p_1)\bar{u}^m_{s_3}(p_3)\gamma^{\nu}v^m_{s_4}(p_4)\\ \times e^{-i(p_1+p_2)\cdot x}e^{i(p_3+p_4)\cdot x'}\int\frac{d^4k}{(2\pi)^4}e^{-ik\cdot(x-x')}\left\{\frac{-ig_{\mu,\nu}}{k^2+i\epsilon}\right\}.\end{aligned} \tag{4.26}$$

The coordinates x and x' are now entirely in the phases and the integrals over these four-dimensional variables give:

$$\begin{aligned}\langle f|\, S\,|i\rangle = ie^2(2\pi)^4 \int \frac{d^4k}{k^2+i\epsilon}\delta^4(p_1+p_2+k)\delta^4(p_3+p_4+k)\\ \times \bar{v}_{s_2}(p_2)\gamma^\mu u_{s_1}(p_1)\bar{u}^m_{s_3}(p_3)\gamma_\mu v^m_{s_4}(p_4).\end{aligned} \tag{4.27}$$

The argument of the first delta function can be rewritten to reflect the presence of the second delta function, and the $\int d^4k$ replaces k^2 in the denominator with $s = (p_1+p_2)^2$. Since this is a positive quantity $s > 4m_\mu^2$ in order for the reaction to be energetically possible, the factor $i\epsilon$ can be set to zero without penalty. This leaves us with the formula:

$$\begin{aligned}\langle f|\, S\,|i\rangle = i(2\pi)^4\delta^4(p_1+p_2-p_3-p_4)\\ \times \left\{\frac{e^2}{s}\bar{v}_{s_2}(p_2)\gamma^\mu u_{s_1}(p_1)\bar{u}^m_{s_3}(p_3)\gamma_\mu v^m_{s_4}(p_4)\right\}.\end{aligned} \tag{4.28}$$

Comparison with Equation (4.3) shows that the expression in the curly brackets is the T matrix element, and it can be substituted directly into the cross-section formula Equation (4.10). Expressions of the form $\bar{u}\gamma^\mu u$ or $\bar{v}\gamma^\mu u$ are 4×4 matrices sandwiched between a row vector on the left and a column vector on the right, and are complex numbers, that commute with each other.

Feynman's rules, which we shall soon introduce, allow the T matrix element to be written down from the diagram, without going through the field theoretical analysis that we just carried out. We went through the complete procedure for a fairly simple case, namely $e^-e^+ \to \mu^-\mu^+$, for at least two reasons. One is to demonstrate that the field theory machinery explained in Chapter 3 can actually be useful, and the other is to appreciate the contribution that Feynman made to our subject.

We still have some work to do to obtain a cross-section. First we square the T matrix, average over initial spins, and sum over final spins:

$$\frac{1}{4}\sum_{s_1,s_2,s_3,s_4} |\langle f|\,T\,|i\rangle|^2 = \frac{e^4}{(4s^2)}\sum_{\text{spins}} \bar{v}_{s_2}(p_2)\gamma^\mu u_{s_1}(p_1)\bar{u}^m_{s_3}(p_3)\gamma_\mu v^m_{s_4}(p_4)$$
$$\times\, \bar{u}_{s_1}(p_1)\gamma^\nu v_{s_2}(p_2)\bar{v}^m_{s_4}(p_4)\gamma_\nu u^m_{s_3}(p_3). \quad (4.29)$$

The electron and muon sums can be performed separately. Group the electron terms together, and write it in terms of the matrix elements in the space of the γ matrices:

$$\begin{aligned}&\frac{1}{4}\sum_{s_1,s_2} \bar{u}_{s_1}(p_1)\gamma^\nu v_{s_2}(p_2)\bar{v}_{s_2}(p_2)\gamma^\mu u_{s_1}(p_1)\\ &\quad = \frac{1}{4}\sum_{\alpha,\beta,\rho,\sigma}\sum_{\text{spins}} \bar{u}_\alpha(p_1)\gamma^\nu_{\alpha,\beta} v_\beta(p_2)\bar{v}_\rho(p_2)\gamma^\mu_{\rho,\sigma} u_\sigma(p_1),\end{aligned} \quad (4.30)$$

where we have suppressed the spin indices on the spinors to avoid clutter. Now we may use the completeness relations:

$$\sum_{s_1} u_\sigma(p_1)\bar{u}_\alpha(p_1) = (\not{p}_1)_{\sigma,\alpha} + m_e\delta_{\sigma,\alpha}, \quad (4.31)$$

and

$$\sum_{s_2} v_\beta(p_2)\bar{v}_\rho(p_2) = (\not{p}_2)_{\beta,\rho} - m_e\delta_{\beta,\rho}. \quad (4.32)$$

The sum over spins and the completeness relations eliminate the spinors from the formula. While it is sometimes desirable to retain the spinors for calculations of polarization phenomena, we shall in most cases use completeness, but introduce spin projection operators into the calculation. Equation (4.30) then becomes a trace, the sum of diagonal elements of a matrix product:

$$\begin{aligned}&\sum_{s_1,s_2} \bar{u}_{s_1}(p_1)\gamma^\nu v_{s_2}(p_2)\bar{v}_{s_2}(p_2)\gamma^\mu u_{s_1}(p_1)\\ &\quad = \mathrm{Tr}\{(\not{p}_1 + m_e)\gamma^\nu(\not{p}_2 - m_e)\gamma^\mu\}\\ &\quad = \mathrm{Tr}(\not{p}_1\gamma^\nu\not{p}_2\gamma^\mu - m_e^2\gamma^\nu\gamma^\mu - m_e\not{p}_1\gamma^\nu\gamma^\mu + m_e\gamma^\nu\not{p}_2\gamma^\mu).\end{aligned} \quad (4.33)$$

It is useful to check that the indices in Equations (4.30)–(4.32) do indeed give Equation (4.33).

4.4. Trace Theorems

Traces of products of γ matrices occur all the time in field theory calculations. The trace is the sum of the diagonal elements of a matrix. The γ matrices are 4×4, and the trace of a 4×4 unit matrix is four. The 2×2 Pauli spin matrices have trace zero, but their squares are one, so $\mathrm{Tr}(\sigma_j^2) = 2$. First a review of some definitions. Dirac's original form:

$$\alpha_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}; \quad \alpha_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix};$$

$$\alpha_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{4.34}$$

and

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{4.35}$$

Each matrix has square one and trace zero, just like the Pauli matrices. They obey the anticommutation rules:

$$\alpha_i \alpha_j + \alpha_j \alpha_i = 2\delta_{i,j}; \quad \beta\alpha_j + \alpha_j \beta = 0. \tag{4.36}$$

The γ matrices are defined as $\gamma^0 = \beta$, $\vec{\gamma} = \beta\vec{\alpha}$, and satisfy the anticommutation rules:

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}, \tag{4.37}$$

where $g^{\mu\nu}$ is the metric tensor, so $(\gamma^0)^2 = 1$, but $(\gamma^j)^2 = -1$. For an arbitrary four vector a^μ that is not a matrix,

$$\not{a} = a^\mu \gamma_\mu = a^\mu \gamma^\nu g_{\mu\nu}. \tag{4.38}$$

A few relations that are not traces are the following:

$$\gamma^{\mu}\gamma_{\mu} = (\gamma^0)^2 - (\vec{\gamma})^2 = 4. \tag{4.39}$$

$$\gamma^{\mu}\not{a}\gamma_{\mu} = \gamma^{\mu}\gamma^{\rho}\gamma_{\mu}a_{\rho} = (-\gamma^{\rho}\gamma^{\mu}\gamma_{\mu} + 2g^{\rho\mu}\gamma_{\mu})a_{\rho} = -4\not{a} + 2\not{a} = -2\not{a}, \tag{4.40}$$

$$\gamma^{\mu}\not{a}\not{b}\gamma_{\mu} = 4a\cdot b, \tag{4.41}$$

$$\gamma^{\mu}\not{a}\not{b}\not{c}\gamma_{\mu} = -2\not{c}\not{b}\not{a}. \tag{4.42}$$

The techniques to prove Equations (4.41) and (4.42) are the same as used in Equation (4.40) — move through the γ matrices using the anticommutation rules so that $\gamma^{\mu}\gamma_{\mu} = 4$.

The trace operation for a product of two matrices is independent of the order, that is $\mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}) = \mathrm{Tr}(\gamma^{\nu}\gamma^{\mu})$, so from the anticommutation rule Equation (4.37) (remember that the metric tensor is a number, depending on μ and ν, times the unit matrix in the space of the γ's) we have

$$\mathrm{Tr}(\not{a}\not{b}) = 4a\cdot b, \quad \text{or} \quad \mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}) = 4g^{\mu\nu}. \tag{4.43}$$

The trace of an odd number of γ matrices vanishes. For the product of four gamma matrices the result is

$$\mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}\gamma^{\rho}\gamma^{\sigma}) = 4(g^{\mu\nu}g^{\rho\sigma} + g^{\mu\sigma}g^{\nu\rho} - g^{\mu\rho}g^{\nu\sigma}). \tag{4.44}$$

The order can be reversed in Equation (4.44), and this inversion is possible in the trace of a larger product of an even number of γ matrices. The trace of a product of six γ matrices can be reduced to five traces of four γ matrices as follows:

$$\begin{aligned}\mathrm{Tr}(\not{a}\not{b}\not{c}\not{d}\not{e}\not{f}) &= (a\cdot b)\mathrm{Tr}(\not{c}\not{d}\not{e}\not{f}) - (a\cdot c)\mathrm{Tr}(\not{b}\not{d}\not{e}\not{f}) + (a\cdot d)\mathrm{Tr}(\not{b}\not{c}\not{e}\not{f})\\ &\quad - (a\cdot e)\mathrm{Tr}(\not{b}\not{c}\not{d}\not{f}) + (a\cdot f)\mathrm{Tr}(\not{b}\not{c}\not{d}\not{e}).\end{aligned} \tag{4.45}$$

The sign alternates, and the dot product with a marches across the other five vectors. This process can be extended to eight or more γ matrices, but it does get tedious.

The special matrix γ^5 introduced in Chapter 2, Equation (2.190), plays an important role in the weak interactions, and trace theorems involving it will be needed in later chapters. We repeat its properties:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3;\ (\gamma^5)^\dagger = \gamma^5;\ (\gamma^5)^2 = 1;\ \mathrm{Tr}(\gamma^5) = 0. \tag{4.46}$$

and

$$\gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0;\quad \gamma^5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{4.47}$$

Under Lorentz transformation $(\bar{\psi}\gamma^5\psi)$ is a pseudoscalar — a scalar under linear transformations, but changing sign for parity and time reversal. The weak interaction employs the axial vector form $(\bar{\psi}\gamma^\mu\gamma^5\psi)$. The non-vanishing trace involving γ^5 has all four γ matrices:

$$\mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^5) = 4i\epsilon^{\mu\nu\rho\sigma}. \tag{4.48}$$

The properties of the four-dimensional antisymmetric tensor $\epsilon^{\mu\nu\rho\sigma}$ were reviewed in Chapter 2, Equations (2.28)–(2.32). The upper index tensor is negative, that is $\epsilon^{0123} = -1$, while $\epsilon_{0123} = +1$. Theorems for contractions of products of the upper and lower index tensors are given in Chapter 2.

4.5. $e^-e^+ \to \mu^-\mu^+$ Cross-Section

We return to Equation (4.33). Obviously the terms linear in the electron mass vanish. The term quadratic in the electron mass is completely negligible, since $s > 4m_\mu^2$ corresponds to a fully relativistic electron with $E/m_e > 200$. However, we shall retain the term, because the muon expression has the same form, and the muon mass may not be negligible. So our task is to evaluate:

$$\begin{aligned} &\mathrm{Tr}(\not{p}_1\gamma^\nu\not{p}_2\gamma^\mu - m_e^2\gamma^\nu\gamma^\mu) \\ &\quad = 4(g^{\alpha\nu}g^{\beta\mu} + g^{\alpha\mu}g^{\nu\beta} - g^{\alpha\beta}g^{\nu\mu})p_{1\alpha}p_{2\beta} - 4m_e^2 g^{\nu\mu} \\ &\quad = 4(p_1^\nu p_2^\mu + p_1^\mu p_2^\nu - (p_1\cdot p_2 + m_e^2)g^{\mu\nu}) = E^{\mu\nu}. \end{aligned} \tag{4.49}$$

This electron tensor $E^{\mu\nu}$ is contracted with the muon tensor $M_{\mu\nu}$ which has the same form to obtain the desired expression (Equation (4.29)):

$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}} |\langle f|\,T\,|i\rangle\,|^2 &= \frac{e^4}{(4s^2)} E^{\mu\nu} M_{\mu\nu} \\ &= \frac{8e^4}{s^2}((p_1 \cdot p_3)(p_2 \cdot p_4) + (p_1 \cdot p_4)(p_2 \cdot p_3) \\ &\quad + m_\mu^2 (p_1 \cdot p_2)). \end{aligned} \tag{4.50}$$

We have dropped the electron mass term, and used the fact that $g^{\mu\nu} g_{\mu\nu} = \mathrm{Tr}(g^2) = 4$. The first two terms are dot products of initial state and final state four vectors, and the third term is $m_\mu^2(s - m_e^2 - m_\mu^2)/2$. The result is obviously Lorentz invariant. Because the electron and positron have exactly the same mass, as do the negative and positive muon, $(p_1 \cdot p_3) = (p_2 \cdot p_4)$, and $(p_1 \cdot p_4) = (p_2 \cdot p_3)$. For $s > 4\,\mathrm{GeV}^2$, in the region of the $3.1\,\mathrm{GeV}$ J/ψ resonance, the final state muons are fully relativistic, and all of the mass terms can safely be neglected. In the center of mass all energies are the same: $E_1 = E_2 = E_3 = E_4$, and $s = 4E^2$. Then the dot products become:

$$p_1 + p_2 = p_3 + p_4, \tag{4.51}$$

$$s = (p_1 + p_2)^2 = 2(p_1 \cdot p_2) = 2(p_3 \cdot p_4), \tag{4.52}$$

$$(p_1 \cdot p_3) = (p_2 \cdot p_4) = \frac{s}{4}(1 - \cos(\theta)) = -t/2, \tag{4.53}$$

$$(p_1 \cdot p_4) = (p_2 \cdot p_3) = \frac{s}{4}(1 + \cos(\theta)) = -u/2, \tag{4.54}$$

where the Mandelstam variables (s, t, u) introduced in Chapter 1 have been evaluated in the high energy limit in the center of mass of the collision — the laboratory of a colliding beam experiment. In this limit, Equation (4.50) can be written:

$$\frac{1}{4}\sum_{\text{spins}} |\langle f|\,T\,|i\rangle\,|^2 = 2e^4 \frac{t^2 + u^2}{s^2}. \tag{4.55}$$

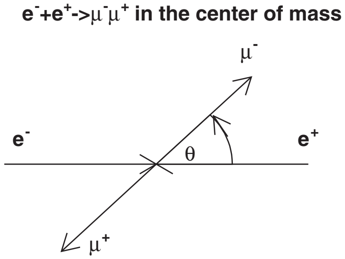


Figure 4.1: Electron–positron annihilation in the center of mass — the lab frame for an e^-e^+ storage ring collider. The polar angle θ between the e^- and μ^- is shown.

The polar angle between the incident electron and the outgoing negative muon is θ, shown in Figure 4.1. The factor s^2 cancels out, and Equation (4.55) becomes:

$$\frac{1}{4}\sum_{\text{spins}} |\langle f|\, T\, |i\rangle\, |^2 = e^4(1+\cos^2(\theta)). \tag{4.56}$$

This expression is dimensionless, as was claimed for the T matrix in a $2 \to 2$ scattering. The s^2 in the denominator comes from the photon propagator, and the fourth power of the momentum in the numerator comes from the completeness sums of the spinors. The initial and final state particles have spin 1/2, so the total spin can be one or zero. Spin zero does not create an intermediate spin one photon. The quantization axis is along the incident beams, for which $J_z = \pm 1$, that is the incident electron and positron are polarized along their momentum vectors. These are called helicity states, for which $\vec{\sigma}\cdot\hat{p} = \pm 1$. Positive helicity is called 'right handed', and negative helicity 'left handed'. If the electron is left handed, the positron is right handed, and vice versa. The outgoing muons are similarly polarized.

The angular dependence $(1+\cos^2(\theta))$ in Equation (4.55) can be explained by appealing to the G-parity problem in Chapter 1, that involves the rotation operator about the y-axis for angular momentum equal to one. We repeat it here:

$$e^{i\theta J_y} = \begin{pmatrix} (1+\cos\theta)/2 & \sin\theta/\sqrt{2} & (1-\cos\theta)/2 \\ -\sin\theta/\sqrt{2} & \cos\theta & \sin\theta/\sqrt{2} \\ (1-\cos\theta)/2 & -\sin\theta/\sqrt{2} & (1+\cos\theta)/2 \end{pmatrix}. \tag{4.57}$$

Operating on $|j=1,m=+1\rangle$ gives:

$$\begin{aligned}
&e^{i\theta J_y}\,|j=1,m=+1\rangle\\
&\quad=(1+\cos\theta)/2\,\big|j=1,m'=+1\big\rangle-\sin\theta/\sqrt{2}\,\big|j=1,m'=0\big\rangle\\
&\qquad+(1-\cos\theta)/2\,\big|j=1,m'=-1\big\rangle. \qquad (4.58)
\end{aligned}$$

The $|j=1,m'=0\rangle$ term does not contribute, and the sum of the squares of the other two terms is $(1+\cos^2(\theta))/2$. Hence the e^- and e^+ have opposite helicities in the initial state, the spins add, and there is no orbital angular momentum.

It is useful to calculate this angular distribution without taking the traces, that is, to use specific forms of the spinors. The procedure is straightforward, if a bit lengthy, but will reward us in an understanding of the roles that the spinors play. It is like taking the back off a watch to see the wheels go around. You do not need to do that if you only want to know what time it is. The initial state matrix element is:

$$\bar{v}^{s_2}(-\vec{p})\gamma^\mu u^{s_1}(\vec{p});\ \vec{p}=E\hat{z};\ E\gg m_e,$$

$$\begin{aligned}
&\mu=0;\ \bar{v}^{s_2}(-\vec{p})\gamma^0 u^{s_1}(\vec{p})=v^{\dagger s_2}(-\vec{p})u^{s_1}(\vec{p})=0,\\
&\mu=3;\ \bar{v}^{s_2}(-\vec{p})\gamma^3 u^{s_1}(\vec{p})=v^{\dagger s_2}(-\vec{p})\alpha_3 u^{s_1}(\vec{p})\\
&\quad=v_a^{\dagger s_2}\sigma_z u_b^{s_1}+v_b^{\dagger s_2}\sigma_z u_a^{s_1}=\begin{pmatrix}0, s_1=s_2\\ 2m_e, s_1=1, s_2=2\\ -2m_e s_1=2, s_2=1\end{pmatrix}. \qquad (4.59)
\end{aligned}$$

The operator γ^0 does not contribute because the two spinors are orthogonal, as noted in Chapter 2. In the high energy limit, where m_e can be neglected, γ^3 does not contribute either. The mass terms correspond to $s_1\neq s_2$, that means the initial state spins along the beam axis cancel. We shall see that $s_1=s_2$ corresponds to $m_s=\pm1$; in other words $s_2=1$ means a negative helicity positron. We have introduced the notation u_a and u_b for the upper and lower two components of the spinor, respectively. We now evaluate γ^1 and the only imaginary matrix γ^2:

$$\begin{aligned}
&\mu=1:\ \bar{v}^{s_2}(-\vec{p})\gamma^1 u^{s_1}(\vec{p})\\
&\quad=v_a^{\dagger s_2}\sigma_x u_b^{s_1}+v_b^{\dagger s_2}\sigma_x u_a^{s_1}=\begin{pmatrix}2E\ s_1=s_2\\ 0\ s_1\neq s_2\end{pmatrix}, \qquad (4.60)
\end{aligned}$$

$$\mu = 2: \ \bar{v}^{s_2}(-\vec{p})\gamma^2 u^{s_1}(\vec{p})$$

$$= v_a^{\dagger s_2}\sigma_y u_b^{s_1} + v_b^{\dagger s_2}\sigma_y u_a^{s_1} = \begin{pmatrix} 2iE \ s_1 = s_2 = 1 \\ -2iE \ s_1 = s_2 = 2 \\ 0 \ s_1 \neq s_2 \end{pmatrix}. \qquad (4.61)$$

The electron tensor then is

$$E^{\mu,\nu} = \frac{1}{4}\sum_{\text{spins}} \bar{v}^{s_2}\gamma^\mu u^{s_1}\bar{u}^{s_1}\gamma^\nu v^{s_2} = 2E^2(\delta_{\mu,1}\delta_{\nu,1} + \delta_{\mu,2}\delta_{\nu,2}). \qquad (4.62)$$

The two Kronecker deltas have equal contributions from $s_1 = s_2 = 1$ and $s_1 = s_2 = 2$, corresponding to $m_s = +1$ and $m_s = -1$, respectively. The cross terms $\sim \gamma^1\gamma^2$ cancel in pairs (an imaginary term plus its complex conjugate). This same formula can be obtained from the Trace Equation (4.49) with $p_2^0 = p_1^0 = E$, and $p_2^3 = -p_1^3 = -E$. Try it!

The weak interaction prefers left handed electrons and right handed positrons. Electrodynamics has no preference, but requires the initial spin to add to ±1 along the beam axis.

The angular dependence appears in the final state muon spinors. Since $M_{\mu\nu}$ couples to $E^{\mu\nu}$, we need only evaluate the γ_1 and γ_2 terms. We choose the scattering plane to be (x, z). Then if $\vec{p}_3$ is the μ^- and $\vec{p}_4$ the μ^+ momentum in the final state, we have:

$$p_{3x} = E\sin(\theta); \quad p_{3z} = E\cos(\theta);$$
$$p_{4x} = -E\sin(\theta); \quad \text{and} \quad p_{4z} = -E\cos(\theta). \qquad (4.63)$$

Evaluation of the matrix elements is straightforward. The spin quantization axis remains along $\hat{z}$, the beam direction. For γ_1:

$$\bar{v}^{s_4}\gamma_1 u^{s_3} = v_a^{\dagger s_4}\sigma_x u_b^{s_3} + v_b^{\dagger s_4}\sigma_x u_a^{s_3}$$

$$= \begin{pmatrix} s_3 = s_4: \ 2E\cos^2(\theta) \\ s_3 = 1, \ s_4 = 2: \ -2E\sin(\theta)\cos(\theta) \\ s_3 = 2, \ s_4 = 1: \ 2E\sin(\theta)\cos(\theta) \end{pmatrix}. \qquad (4.64)$$

And for γ_2:

$$\bar{v}^{s_4}\gamma_2 u^{s_3} = v_a^{\dagger s_4}\sigma_y u_b^{s_3} + v_b^{\dagger s_4}\sigma_y u_a^{s_3} = \begin{pmatrix} s_3 = s_4 = 1:\ 2iE \\ s_3 = s_4 = 2:\ -2iE \\ s_3 \neq s_4:\ 0 \end{pmatrix}. \tag{4.65}$$

The muon tensor is

$$M_{\mu\nu} = 8E^2(\cos^2(\theta)\delta_{\mu,1}\delta_{\nu,1} + \delta_{\mu,2}\delta_{\nu,2}). \tag{4.66}$$

The contraction of the two tensors gives:

$$E^{\mu\nu}M_{\mu\nu} = 16E^4(\cos^2(\theta) + 1). \tag{4.67}$$

The formula is in agreement with the result from using the trace theorems. Convert $16E^4 = s^2$ for comparison. This exercise puts the assertion that it is $v^{(1)}(p_2)$ that has negative helicity on firm ground. The e^-e^+ pair forms $s = 1, m_s = \pm 1$ along the beam to couple to the spin one intermediate photon. The $\mu^-\mu^+$ pair in the final state also has $m_s = \pm 1$ along their common direction of motion, although that is not obvious from this analysis, because the spinors are still quantized along the incident beam axis. At $\theta = \pi/2$ for example, the formula is just one, which corresponds to equal probability for either muon to have $\sigma_z = +1$ or $\sigma_z = -1$, consistent with the spin along the x axis. A spin projection operator will be introduced in Section 4.10 that takes into account the direction of motion in the final state.

We now have all the parts to calculate the cross-section from Equation (4.10). It is useful to evaluate everything except the T matrix element for our special case of a two-body-to-two-body highly relativistic reaction in the center of mass. So, let us rewrite Equation (4.10) as follows:

$$d\sigma_{i\to f} = \frac{(2\pi)^4}{(2\pi)^6}\frac{1}{2s}\delta(\sqrt{s} - E_3 - E_4)\delta(\vec{p}_3 + \vec{p}_4)\frac{d^3p_3}{2E_3}\frac{d^3p_4}{2E_4}|\langle f|\,T\,|i\rangle|^2. \tag{4.68}$$

The factor $2s$ in the denominator comes from the flux, and the δ functions have been recast to reflect the fact that we are in the center of mass, where the total momentum is zero. Integration over d^3p_4 eliminates the momentum δ function, and sets $\vec{p}_3 = -\vec{p}_4$, so the energy delta function becomes $\delta(\sqrt{s}-2E_4)$. The factor of two in the argument appears in the denominator $\int p_3^2 dp_3 d\Omega \delta(\sqrt{s}-2E_4) = d\Omega \int p_3 E_3 dE_3 \delta(\sqrt{s}/2 - E_4)/2 = p_3 E_3 d\Omega/2$. These manipulations give us

$$\frac{d\sigma_{i\to f}}{d\Omega} = \frac{|\langle f|\, T\,|i\rangle\,|^2}{64\pi^2 s}. \tag{4.69}$$

This formula is very useful, as it avoids the necessity of evaluating phase space and flux terms every time, provided that the conditions are satisfied, namely, two body to two body in the center of mass, and $\sqrt{s} \gg m$, all masses neglected. Phase space will return for a three body final state, and even though Compton scattering is two to two, we will evaluate the cross-section in the laboratory frame, where Compton first did the experiment. The initial electron is at rest, and the electron mass will not be neglected, so for Compton scattering we will have to start over with the kinematics.

The fine structure constant can be used to replace $e^4 = 16\pi^2\alpha^2$. Substituting Equation (4.55) for the unpolarized T matrix gives the differential cross-section:

$$\frac{d\sigma(e^-e^+ \to \mu^-\mu^+)}{d\Omega} = \frac{\alpha^2}{4s}(1+\cos^2\theta). \tag{4.70}$$

The polar angle $0 \le \theta \le \pi$ of the μ^- is measured with respect to the incoming electron momentum; the unpolarized cross-section is independent of the azimuthal angle in the plane perpendicular to the beam $0 \le \phi \le 2\pi$. The initial and final state particles are distinguishable, and the cross-section is front-back symmetric — the initial direction of the flow of negative charge (the electron) is not retained by the intermediate photon. The negative muon is just as likely to go backwards as forwards.

Integration over the solid angle gives $16\pi/3$, so one unit of R is:

$$\sigma = \frac{4\pi\alpha^2}{3s} = 8.7\,\text{nb} = 8.7 \times 10^{-33}\text{cm}^2 \text{ at } s = 10\,\text{GeV}^2. \tag{4.71}$$

As mentioned in Chapter 1, this cross-section is used as a unit for the measurement of other cross-sections in e^-e^+ annihilation. Quoting $\sigma(e^-e^+ \to c\bar{c})$ in units of R for example takes out the general 1/s decrease in the cross-section with increasing energy. The form $\sigma \sim \alpha^2/s$ could have been argued on dimensional grounds, so all of this work only got us the $4\pi/3$.

4.6. $e^-\mu^- \to e^-\mu^-$

This process looks very much like $e^-e^+ \to \mu^-\mu^+$. Simply move the positron to the other side, change it into its antiparticle the electron, and do the same in reverse for the μ^+. This is an example of what is known as crossing symmetry — switching particles from initial to final state or vice versa, and changing them into their antiparticles. The diagrams are closely related, except that $e^-\mu^-$ scattering has no antiparticles, so no spinors $v_s(p)$. Experimentally $e^-\mu^-$ scattering is difficult to study. There are no $e^-\mu^-$ colliding beams, and there are no muon targets. There are muon beams, and electron targets, so in principle the process is accessible, but I am not aware of any relevant experimental data. Write down the two amplitudes based on their diagrams:

- for $e^-e^+ \to \mu^-\mu^+$

$$-e^2\bar{v}(p_2)\gamma^\mu u(p_1)\frac{-ig_{\mu\nu}}{s}\bar{u}(p_3)\gamma^\nu v(p_4);\; s = (p_1+p_2)^2 \tag{4.72}$$

- and for $e^-\mu^- \to e^-\mu^-$

$$-e^2\bar{u}(k_3)\gamma^\mu u(k_1)\frac{-ig_{\mu\nu}}{t}\bar{u}(k_4)\gamma^\nu u(k_2);\; t = (k_1-k_3)^2. \tag{4.73}$$

Here, $\bar{u}(k_4)$ and $u(k_2)$ are muon spinors, and we have suppressed the spin indices throughout. The four momenta are exchanged as follows:

$$p_1 \to k_1;\ p_2 \to -k_3;\ p_3 \to k_4;\ p_4 \to -k_2. \tag{4.74}$$

For the Mandelstam variables:

$$s \to t;\ t \to u;\ u \to s. \tag{4.75}$$

Note that the particles that are switched have their four momenta reversed, because they have changed from initial to final state, or vice versa. Then from Equation (4.50) for $e^-e^+ \to \mu^-\mu^+$ we can write down the corresponding expression for $e^-\mu^- \to e^-\mu^-$:

$$\frac{1}{4}\sum_{\text{spins}} |\langle f|\,T\,|i\rangle|^2 = \frac{8e^4}{t^2}((k_1\cdot k_4)(k_2\cdot k_3) + (k_1\cdot k_2)(k_3\cdot k_4) - m_\mu^2(k_1\cdot k_3)). \tag{4.76}$$

The four momenta have been shuffled around, and the mass term has a minus sign, but perhaps the most important difference comes from the photon propagator. Instead of the well behaved $1/s^2$ we have $1/t^2$, and that diverges for $\theta = 0$. The Coulomb pole is the characteristic of Rutherford scattering discussed in Chapter 3. In terms of the Mandelstam variables $s = (k_1+k_2)^2;\ t = (k_1-k_3)^2;\ u = (k_1-k_4)^2;\ s+t+u = 2m_\mu^2$, we have:

$$\frac{1}{4}\sum_{\text{spins}} |\langle f|\,T\,|i\rangle|^2 = \frac{2e^4}{t^2}(u^2+s^2-2m_\mu^4). \tag{4.77}$$

We may leave it in this compact form, valid in any Lorentz frame provided that the electron mass can be neglected. In the high energy limit where the muon mass may also be neglected, the expression simplifies in the center of mass of the $e^-\mu^-$ collision:

$$\frac{1}{4}\sum_{\text{spins}} |\langle f|\,T\,|i\rangle|^2 = 2e^4\left\{\frac{u^2+s^2}{t^2}\right\} = 2e^4\left\{\frac{1+\cos^4(\theta/2)}{\sin^4(\theta/2)}\right\}, \tag{4.78}$$

where θ is the polar angle between the incoming and outgoing electron. This formula differs from that of Rutherford by the addition of a relativistic term in the numerator, $\cos^4(\theta/2)$. Otherwise,

$e^-\mu^-$ scattering would be just like α particles on gold. (In the nonrelativistic limit the numerator is a constant, and in the denominator $t \to -2p^2(1 - \cos\theta)$ in both limits!)

In solving this second problem based on the first, we have learned something about how to interpret the diagrams. They are more than just pictures. Substitution of Equation (4.78) into Equation (4.69) gives the differential cross-section:

$$\frac{d\sigma(e^-\mu^- \to e^-\mu^-)}{d\Omega} = \frac{\alpha^2}{2s}\left\{\frac{1+\cos^4(\theta/2)}{\sin^4(\theta/2)}\right\}. \tag{4.79}$$

4.7. Feynman Diagrams and Feynman Rules

The diagrams are just as important as the rules. The two go together. Given the reaction of interest, the contributing diagrams have to be accurately drawn, because the assembly of the T matrix element is specified by the diagram. This procedure is a great simplification of the field theoretic technique that we outlined in the previous section. It gives you an amplitude $-i\langle f|\,T\,|i\rangle$ by inspection. The rules for QED are summarized in Figure 4.2. Rules for weak interactions will be introduced later. From the picture form a scalar product (a row vector times a matrix times a column vector) at one vertex $-ie\bar{u}(p_3)\gamma^\mu u(p_1)$. Do the same thing at the other vertex:$-ie\bar{u}(p_4)\gamma^\nu u(p_2)$. Now hook them together with the photon propagator $-ig_{\mu\nu}/q^2$, where $q = p_1 - p_3$, and you have one amplitude for Möller scattering. You still have to write down the crossed amplitude, square the difference, sum over the spins, take all of the necessary traces, and deal with the final state phase space to get the cross-section. For some elementary problems this can be fairly difficult. So, even with Feynman's huge contributions, problems in quantum electrodynamics can be a challenge.

The diagrams have lines and vertices. Four momentum is conserved at each vertex. The basic Hamiltonian density is $\mathcal{H} = -e\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu(x)$. The Feynman rules work in momentum space, and we have seen in the examples just considered how configuration space functions and integrals become momentum space expressions.

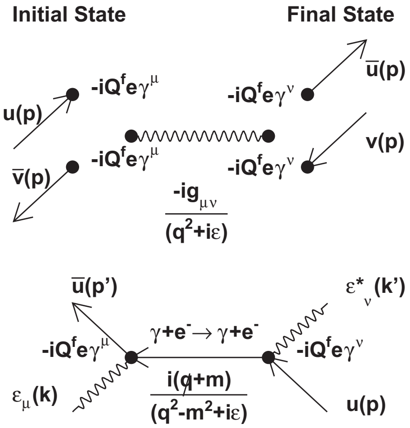

Initial State
Final State
u(p)
$\bar{v}$(p)
-iQ^{f}eγ^μ
-iQ^{f}eγ^μ
-iQ^{f}eγ^ν
-iQ^{f}eγ^ν
$\bar{u}$(p)
v(p)
-ig$_{\mu\nu}$ / (q^2+iε)
$\bar{u}$(p')
ε^*_ν(k')
γ+e$^-$→ γ+e$^-$
-iQ^{f}eγ^μ
-iQ^{f}eγ^ν
i($\not{q}$+m) / (q^2-m^2+iε)
ε_μ(k)
u(p)

Figure 4.2: Feynman Rules for charged fermions — e's, μ's, τ's and quarks — and photons. Q^f is the fermion charge, and e is positive. For an electron $Q^f = -1$, and for a u quark $Q^f = +2/3$. The correct four momentum must be inserted in the spinors. Antifermion lines are directed opposite to the momentum. Real photons have momentum k, and the four momentum in the propagators is $q = p_{\text{in}} - p_{\text{out}}$. Begin with the left vertex, sandwich $\bar{u}(p_3)(-ie\gamma^\mu)u(p_1)$. Couple the other vertex with the photon propagator $-ig_{\mu\nu}/q^2$. The sandwich with three $-i$ gives $-i\langle f|\,T\,|i\rangle$. One of the diagrams for Compton scattering is shown on the bottom. The electron propagator involves $q = k - p'$. Interchanging two electron lines or exchanging an outgoing positron for an incoming electron changes the sign of the amplitude. Interchanging two photon lines gives a plus sign.

Negative leptons in the initial state are straight arrows, pointing towards the vertex. Each incoming negative lepton gets a spinor $u_s(p)$. Positive leptons (antiparticles) in the initial state are straight arrows pointing away from the vertex. Each incoming positive lepton gets a spinor $\bar{v}_s(p)$. A real photon in the initial state, as in Compton scattering $\gamma + e^- \to \gamma + e^-$ is a wiggly line terminating on the vertex. Each incoming real photon gets a polarization four vector ϵ^μ. A real photon cannot be emitted from an electron at a single vertex and conserve four momentum. Hence the physical processes which we shall consider require at least two vertices, and two Hamiltonians in the perturbation expansion of the S matrix. Each vertex gets a factor $-ie\gamma^\mu$. In our units the elementary charge $e = 0.3$. If two vertices can describe the physical process, and four vertices is a radiative correction, the four-vertex amplitude is about 10% of the main term. It is this convergence that makes quantum electrodynamics a successful theory. Increasing the complexity leads to more diagrams, and for a convergent theory the coupling parameter must be small enough so that the net effect of the extra terms is less than one.

If a virtual photon couples the two vertices, then a factor $-ig_{\mu\nu}/s$ for s channel exchange, and $-ig_{\mu\nu}/t$ for t channel exchange is inserted. For Compton scattering the intermediate exchanged particle is an off mass shell electron, requiring the electron propagator $i(\not{p} + m)/(p^2 - m^2 + i\epsilon)$, where ϵ can be ignored, as $p^2 \neq m^2$. For the particles in the final state, a real photon gets a curly line and a factor ϵ^*_ν, an electron is straight arrow pointing away from the vertex, $\bar{u}_s(p)$, and a positron is a straight arrow pointing towards the vertex, $v_s(p)$. The final result of assembling all of these terms is $-i\langle f|\, T\, |i\rangle$.

4.8. Bhabha Scattering $e^-e^+ \to e^-e^+$

Bhabha scattering involves two diagrams shown in Figure 4.3. One is the same as $e^-e^+ \to \mu^-\mu^+$, with an e^-e^+ final state, and the other is similar to $e^-\mu^- \to e^-\mu^-$, if the muon is replaced by a positron. There are the s channel annihilation diagram and the t channel exchange diagram. Suppressing the spin indices we have from the

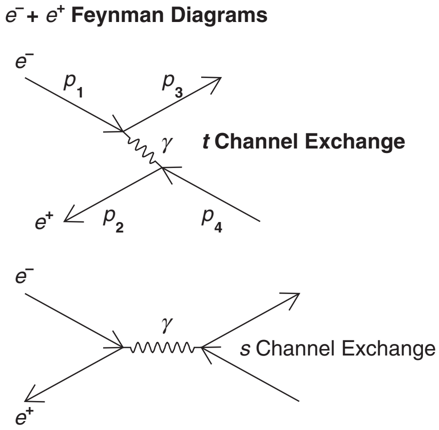


Figure 4.3: s and t channel diagrams for Bhabha scattering. Only the s channel diagram contributes to $e^-e^+ \to \mu^-\mu^+$.

rules: From the s channel diagram

$$-e^2\bar{v}(p_2)\gamma^\mu u(p_1)\frac{-ig_{\mu\nu}}{(p_1+p_2)^2}\bar{u}(p_3)\gamma^\nu v(p_4), \tag{4.80}$$

and from the t channel diagram

$$-e^2\bar{u}(p_3)\gamma^\lambda u(p_1)\frac{-ig_{\lambda\sigma}}{(p_1-p_3)^2}\bar{v}(p_2)\gamma^\sigma v(p_4). \tag{4.81}$$

In the s channel diagram, electrons are coupled to positrons in both the initial and final states, while in the t channel diagram electrons are coupled to electrons and positrons to positrons. Equations (4.72) and (4.80) are identical in form, except the final state is muons in the first and electrons in the second. In both cases, the trace reduces to a product of two independent traces, as can be checked by writing out the expression in components. Equations (4.73) and (4.81) look different, because (4.73) has no antiparticles. The cross-section is proportional to:

$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}}|\langle f|\,S\,|i\rangle|^2 = \frac{1}{4}\sum_{\text{spins}} &|(s\text{ channel amplitude})\\ &-(t\text{ channel amplitude})|^2,\end{aligned} \tag{4.82}$$

and therefore involves two magnitude squares that can be obtained from work that we have already done, plus an interference term that is a new animal. Our first two cross-sections were cleverly chosen to avoid interference terms, but our luck has run out. A minus sign has appeared in Equation (4.82). This is a subtle business, that will occur often when we have to evaluate several diagrams. In this case there is one fermion exchanged between the two diagrams: $p_2 \to -p_3$, and that flips the sign. We will work in the limit $m_e \to 0$. We can then write down the magnitude squared terms directly:

$$\frac{1}{4}\sum_{\text{spins}}|A(s)|^2 = 2e^4\frac{t^2+u^2}{s^2};\ \frac{1}{4}\sum_{\text{spins}}|A(t)|^2 = 2e^4\frac{s^2+u^2}{t^2}. \tag{4.83}$$

The second expression was obtained from the first by the substitutions: $s \to t$, $t \to s$, and $u \to u$. The cross term is

$$-\frac{1}{4}\sum_{\text{spins}}(A^\dagger(s)A(t) + A^\dagger(t)A(s)) = -\frac{1}{2}\sum_{\text{spins}} A^\dagger(t)A(s). \tag{4.84}$$

The two traces are identical, so we only have to do one of them. The expression does not factor into two tensors. This is often a property of an interference term. We are faced with the trace of eight γ matrices, but we shall see that it is not so bad. One of the terms multiplied by two looks like this:

$$\begin{aligned} -\frac{1}{2}\sum_{\text{spins}} A^\dagger(t)A(s) = -\frac{e^4}{2st}\sum_{\text{spins}} \big\{&\bar{v}(p_4)\gamma^\sigma v(p_2)\bar{u}(p_1)\gamma_\sigma u(p_3)\bar{v}(p_2) \\ &\times \gamma^\mu u(p_1)\bar{u}(p_3)\gamma_\mu v(p_4)\big\}. \end{aligned} \tag{4.85}$$

This reduces to the trace equation:

$$= -\frac{e^4}{2st}\mathrm{Tr}\big\{\not{p}_4\gamma^\sigma\not{p}_2\gamma^\mu\not{p}_1\gamma_\sigma\not{p}_3\gamma_\mu\big\}. \tag{4.86}$$

If you are uncertain about the order of the operators in a long trace, write it out in component form, and assemble the matrices in correct order. If you run out of Greek letters for the indices, it is permissible to use Roman letters instead. The trace is invariant to the exchange of any pair of matrices. For example,

$$\mathrm{Tr}\big\{\not{p}_4\gamma^\sigma\not{p}_2\gamma^\mu\not{p}_1\gamma_\sigma\not{p}_3\gamma_\mu\big\} = \mathrm{Tr}\big\{\not{p}_2\gamma^\mu\not{p}_1\gamma_\sigma\not{p}_3\gamma_\mu\not{p}_4\gamma^\sigma\big\}.$$

The same index should appear on the far left and the far right (it is a trace!). This expression reduces to the trace of four gamma matrices by using the $\gamma^\mu\not{a}\not{b}\gamma_\mu = 4a \cdot b$ and similar theorems. Thus,

$$\begin{aligned} -\frac{1}{2}\sum_{\text{spins}} A^\dagger(t)A(s) &= +\frac{e^4}{st}\mathrm{Tr}\big\{\not{p}_4\not{p}_1\gamma^\mu\not{p}_2\not{p}_3\gamma_\mu\big\} \\ &= \frac{16e^4}{st}(p_1 \cdot p_4)(p_2 \cdot p_3) = 4e^4\frac{u^2}{st}. \end{aligned} \tag{4.87}$$

For the complex conjugate term, the momentum labels for p_1 and p_2 are interchanged, as are p_3 and p_4, and that leaves the answer unchanged. We have already doubled the answer, so we may assemble all of the parts:

$$\frac{1}{4}\sum_{\text{spins}}|A(t)-A(s)|^2 = 2e^4\left\{\frac{s^2+u^2}{t^2}+\frac{t^2+u^2}{s^2}+2\frac{u^2}{st}\right\}. \tag{4.88}$$

The Mandelstam variables greatly simplify the formula. Since $s+t+u=0$, there are many equivalent variations of Equation (4.88). It can be converted into the polar angle θ:

$$\frac{1}{4}\sum_{\text{spins}}|A(t)-A(s)|^2$$
$$= 4e^4\left\{\frac{1+\cos^4(\theta/2)}{2\sin^4(\theta/2)}+\frac{(1+\cos^2(\theta))}{4}-\frac{\cos^4(\theta/2)}{\sin^2(\theta/2)}\right\}. \tag{4.89}$$

Substitution of the T matrix element into Equation (4.69) gives the cross-section:

$$\frac{d\sigma(e^-e^+\to e^-e^+)}{d\Omega}$$
$$= \frac{\alpha^2}{s}\left\{\frac{1+\cos^4(\theta/2)}{2\sin^4(\theta/2)}+\frac{(1+\cos^2(\theta))}{4}-\frac{\cos^4(\theta/2)}{\sin^2(\theta/2)}\right\}. \tag{4.90}$$

This formula is given by Nachtmann on page 135 of his text [2]. The Coulomb singularity is exhibited in the first term, $|A(t)|^2$, and in the third interference term. The middle term comes from the annihilation diagram.

Bhabha scattering has been extensively studied a e^-e^+ storage ring colliders, and the angular dependence of Equation (4.89) has been confirmed. Small angle Bhabha's are used as a luminosity monitor.

4.9. $e^-e^-\to e^-e^-$ Møller Scattering

Here, we are dealing with identical fermions, and the cross-section must be front–back symmetric. Classically one would expect a cross-section to have a Coulomb pole at both ends:

$$\sigma_{\text{classic}} \sim \frac{1}{\sin^4(\theta/2)}+\frac{1}{\cos^4(\theta/2)}. \tag{4.91}$$

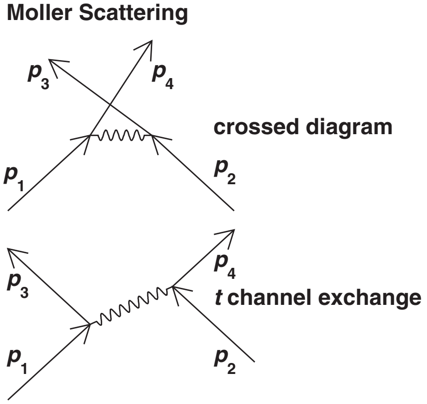


Figure 4.4: Feynman diagrams for Møller scattering. There is an interference term between the two amplitudes, labeled $A(t)$ and $A(u)$ in the text.

The electrons are identical particles, and you do not know which electron winds up in the forward direction, and which in the backward direction. Quantum mechanics forbids the observation of the trajectory. Integration of the differential cross-section has the range $0 \leq \theta \leq \pi/2$, with the usual singularity at $\theta = 0$. Quantum electrodynamics will reproduce Equation (4.91), with the addition of an interference term between forward and backward amplitudes, and some relativistic factors in the numerators. Figure 4.4 shows the two Feynman diagrams. Here, for the first time we meet a crossed diagram where instead of $1 \to 3$ and $2 \to 4$, which we will call $A(t)$, we have $1 \to 4$ and $2 \to 3$, called $A(u)$. Since they are all electrons, no one can tell the difference. As was the case with Bhabha scattering, the fermion exchange in the two diagrams merits a minus sign between them. Thus,

$$\frac{1}{4}\sum_{\text{spins}} |\langle f|\, T\,|i\rangle\,|^2 = \frac{1}{4}\sum_{\text{spins}} |A(t) - A(u)|^2. \tag{4.92}$$

From the diagrams we have:

$$A(t) = -e^2 \bar{u}(p_3)\gamma^\lambda u(p_1) \left\{\frac{-ig_{\lambda\mu}}{t}\right\} \bar{u}(p_4)\gamma^\mu u(p_2); \tag{4.93}$$

and

$$A(u) = -e^2 \bar{u}(p_4)\gamma^\nu u(p_1)\left\{\frac{-ig_{\nu\sigma}}{u}\right\}\bar{u}(p_3)\gamma^\sigma u(p_2). \tag{4.94}$$

For $A(t) \to A(u)$, $p_3 \to p_4$, $t = (p_1 - p_3)^2$, and $u = (p_1 - p_4)^2$. The magnitude squared $|A(t)|^2$ spin summed and averaged is

$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}}|A(t)|^2 = \frac{e^4}{4t^2}\sum_{\text{spins}} &\bar{u}(p_3)\gamma^\lambda u(p_1)\bar{u}(p_4)\gamma_\lambda u(p_2)\\ &\times \bar{u}(p_2)\gamma^\mu u(p_4)\bar{u}(p_1)\gamma_\mu u(p_3);\end{aligned} \tag{4.95}$$

and for $|A(u)|^2$:

$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}}|A(u)|^2 = \frac{e^4}{4u^2}\sum_{\text{spins}} &\bar{u}(p_4)\gamma^\sigma u(p_1)\bar{u}(p_3)\gamma_\sigma u(p_2)\bar{u}(p_2)\\ &\times \gamma^\nu u(p_3)\bar{u}(p_1)\gamma_\nu u(p_4).\end{aligned} \tag{4.96}$$

The expressions are the same with p_3 and p_4 interchanged, and each factors into two traces. For $|A(t)|^2$:

$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}}|A(t)|^2 &= \frac{e^4}{4t^2}\mathrm{Tr}(\gamma_\mu \not{p}_3\gamma^\lambda \not{p}_1)\mathrm{Tr}(\not{p}_4\gamma_\lambda \not{p}_2\gamma^\mu)\\ &= \frac{4e^4}{t^2}(p_{3\mu}p_1^\lambda + p_{1\mu}p_3^\lambda - g_\mu^\lambda(p_1\cdot p_3))\\ &\quad\times(p_{4\lambda}p_2^\mu + p_{2\lambda}p_4^\mu - g_\lambda^\mu(p_2\cdot p_4))\\ &= \frac{8e^4}{t^2}((p_1\cdot p_4)(p_2\cdot p_3) + (p_1\cdot p_2)(p_3\cdot p_4)).\end{aligned} \tag{4.97}$$

We have written the metric tensor g_λ^μ for symmetry in the formula, but it is the same object with upper, lower, or mixed indices. The summed product of the two metric tensors is $\mathrm{Tr}(g^2) = 4$. In terms of the Mandelstam variables this expression is

$$\frac{1}{4}\sum_{\text{spins}}|A(t)|^2 = \frac{2e^4}{t^2}(s^2 + u^2), \tag{4.98}$$

and

$$\frac{1}{4}\sum_{\text{spins}} |A(u)|^2 = \frac{2e^4}{u^2}(s^2 + t^2). \tag{4.99}$$

We still have to calculate the interference term. As usual, it does not factor into two tensors. It suffices to evaluate $A^\dagger(u)A(t)$ and double the result:

$$\begin{aligned}
&\frac{-e^4}{2tu}\sum_{\text{spins}} \bar{u}(p_1)\gamma^\sigma u(p_4)\bar{u}(p_2)\gamma_\sigma u(p_3)\bar{u}(p_3)\gamma^\lambda u(p_1)\bar{u}(p_4)\gamma_\lambda u(p_2) \\
&\quad = \frac{-e^4}{2tu}\mathrm{Tr}\left\{\gamma^\lambda \not{p}_1 \gamma^\sigma \not{p}_4 \gamma_\lambda \not{p}_2 \gamma_\sigma \not{p}_3\right\} \\
&\quad = \frac{16e^4}{tu}((p_3 \cdot p_4)(p_1 \cdot p_2) = \frac{4e^4 s^2}{tu}.
\end{aligned} \tag{4.100}$$

We now have all the parts for Møller scattering:

$$\frac{1}{4}\sum_{\text{spins}} |A(t) - A(u)|^2 = 2e^4 \left\{\frac{s^2 + u^2}{t^2} + \frac{s^2 + t^2}{u^2} + \frac{2s^2}{tu}\right\}. \tag{4.101}$$

Substitution into Equation (4.69) gives the cross-section in the center of mass:

$$\begin{aligned}
\frac{d\sigma(e^-e^- \to e^-e^-)}{d\Omega} = \frac{\alpha^2}{2s}\Bigg\{&\frac{1 + \cos^4(\theta/2)}{\sin^4(\theta/2)} + \frac{1 + \sin^4(\theta/2)}{\cos^4(\theta/2} \\
&+ \frac{2}{\sin^2(\theta/2)\cos^2(\theta/2)}\Bigg\}.
\end{aligned} \tag{4.102}$$

So, to go from Bhabha to Møller scattering, simply interchange s and u. This remarkable feature of the Mandelstam variables, repeated in many different formulas, is the reason for their wide application in relativistic theory. This was not recognized until many years after Mandelstam's introduction in a 1958 paper on pion–nucleon dispersion relations [3]. Current theoretical calculations, and texts, universally use these variables.

Any of these cross-sections can be expressed as $d\sigma/dt$, since $t = -s(1 - \cos\theta)/2$, $dt = (s d\cos\theta)/2 = s d\Omega/(4\pi)$. So, for Møller scattering we may write:

$$\frac{d\sigma(e^-e^- \to e^-e^-)}{dt} = \frac{2\pi\alpha^2}{s^2}\left\{\frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2} + \frac{2s^2}{ut}\right\}. \quad (4.103)$$

4.10. Spin Effects

The Pauli spin matrices play an integral part in our representation of the 4×4 γ matrices that appear in the Dirac equation for a free spin 1/2 electron. It is therefore natural to expect that the spin of the electron and positron play a role in quantum electrodynamics. Spin effects are especially important in weak interactions, the subject of Chapter 6, where parity violation leads to polarized particles. Electrodynamics conserves parity, but there are still interesting spin effects that have been mentioned above in our discussion of the angular distribution in $e^-e^+ \to \mu^-\mu^+$. The vector interaction requires the two spin 1/2 fermions to add to spin 1, not spin zero. We have demonstrated this for the reaction $e^-e^+ \to \mu^-\mu^+$ by working explicitly with the spinors themselves. A more general technique is to perform the spin sums, that lead to traces, and insert spin projection operators into the trace expression.

The spin direction of the electron is specified in the particle center of mass, where $\vec{p} = 0$, and $E = m$. This is a three-dimensional frame, in which we choose the time component of the spin direction four vector to be zero. This four vector is spacelike. In the rest frame:

$$s = (0, \hat{n});\ \ s \cdot s = -1; \quad (4.104)$$

where

$$\hat{n} = \sin\theta\cos\phi\hat{x} + \sin\theta\sin\phi\hat{y} + \cos\theta\hat{z}, \quad \text{and}$$
$$\vec{\sigma}\cdot\hat{n} = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}. \quad (4.105)$$

By the spin in the direction of $\hat{n}$ we mean that the $\langle\vec{\sigma}\cdot\hat{n}\rangle = +1$, or the spin eigenvector is

$$\psi_+ = a\,|\uparrow\rangle + b\,|\downarrow\rangle = \begin{pmatrix} \cos(\theta/2)e^{-i\phi/2} \\ \sin(\theta/2)e^{i\phi/2} \end{pmatrix}. \quad (4.106)$$

We now perform a Lorentz boost along the z-axis. Components of $\hat{n}$ perpendicular to z, in the (x, y) plane, are Lorentz invariant. In what follows, we assume that $\hat{n} = \hat{z}$, that is the electron is longitudinally polarized. The four vector for the spin direction after Lorentz transformation becomes:

$$s = (\beta\gamma, 0, 0, \gamma) = (p/m, 0, 0, E/m), \tag{4.107}$$

and of course $s \cdot s = -1$ is still true. The energy is the space component, and the momentum is the time component — just backwards. The appropriate projection operator for the electron spinor with positive helicity is

$$\Lambda_+ = \frac{1 + \gamma^5 \not{s}}{2} = \frac{1}{2}\begin{pmatrix} 1 + E/m & 0 & -p/m & 0 \\ 0 & 1 - E/m & 0 & -p/m \\ p/m & 0 & 1 - E/m & 0 \\ 0 & p/m & 0 & 1 + E/m \end{pmatrix};$$

$$u^{(1)}(p) = \sqrt{E + m}\begin{pmatrix} 1 \\ 0 \\ p/(E + m) \\ 0 \end{pmatrix}. \tag{4.108}$$

Matrix multiplication shows that $\Lambda_+ u^{(1)}(p) = u^{(1)}(p)$. Similarly $\Lambda_+ u^{(2)}(p) = 0$. Because $s \cdot p = 0$, $\left[\gamma^5 \not{s}, \not{p}\right] = 0$. Specifically:

$$\begin{aligned} \left[\gamma^5 \not{s}, \not{p}\right] &= \gamma^5 \not{s}\not{p} - \not{p}\gamma^5 \not{s} = \gamma^5(\not{s}\not{p} + \not{p}\not{s}) = \gamma^5(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)s_\mu p_\nu \\ &= \gamma^5(2s \cdot p) = 0. \end{aligned} \tag{4.109}$$

Therefore, $\not{p}$ and $\gamma^5 \not{s}$ can have simultaneous eigenvalues, namely m and $+1$, respectively. With $\Lambda_- = \frac{1-\gamma^5 \not{s}}{2}$ we have two projection operators with the usual characteristics, namely,

$$\Lambda_+^2 = \Lambda_+; \Lambda_-^2 = \Lambda_-; \Lambda_+\Lambda_- = 0; \Lambda_+ + \Lambda_- = 1. \tag{4.110}$$

A Lorentz boost from the positron rest frame gives the spinor $v^{(1)}(p)$ (E is positive in the Lorentz transformation, and the positron is going in the opposite direction to the electron):

$$v^{(1)}(p) = \sqrt{E+m}\begin{pmatrix} 0 \\ p/(E+m) \\ 0 \\ 1 \end{pmatrix}, \tag{4.111}$$

and for the positron, because $\vec{p} = -p\hat{z}$, the spin projection operator is

$$\Lambda_+ = \frac{1}{2}\begin{pmatrix} 1+E/m & 0 & p/m & 0 \\ 0 & 1-E/m & 0 & p/m \\ -p/m & 0 & 1-E/m & 0 \\ 0 & -p/m & 0 & 1+E/m \end{pmatrix};$$

$$\Lambda_+ v^{(1)}(p) = v^{(1)}(p). \tag{4.112}$$

However, the positron has negative helicity, that is the spin is pointing along $+\hat{z}$, opposite to its direction of motion, as we have seen in our spinor analysis of the $e^-e^+ \to \mu^-\mu^+$. So in an e^-e^+ head on collision, with the electron traveling in the $+\hat{z}$ direction, $(1+\gamma^5\not{s})/2$ gives total magnetic quantum number $m = +1/2 + 1/2 = +1$ along the beam. The completeness relation can be modified:

$$u^{(s=1)}(p)\bar{u}^{(s=1)}(p) = (\not{p}+m)(1+\gamma^5\not{s})/2; \; \sum_{s=1,2} u^{(s)}(p)\bar{u}^{(s)}(p)$$

$$= (\not{p}+m)(\Lambda_+ + \Lambda_-) = \not{p}+m. \tag{4.113}$$

So, you may perform the spin sum and retain only $u^{(1)}\bar{u}^{(1)}$ if you add in the projection operator Λ_+. The calculation of the electron tensor using the spinors, Equation (4.62), can also be calculated by evaluating the two traces:

$$E^{\mu\nu}_{RL} = \frac{1}{4}\sum_{\text{spins}} \bar{v}^{s_2}\gamma^\mu u^{s_1}\bar{u}^{s_1}\gamma^\nu v^{s_2} = \frac{1}{8}(\mathrm{Tr}(\not{p}_2\gamma^\mu(\not{p}_1+m)(1+\gamma^5\not{s}_1)\gamma^\nu). \tag{4.114}$$

The subscript RL refers to the electron and positron helicity, respectively. The electron mass has to be retained, even though we are working in the relativistic limit $E \gg m_e$, because the electron mass appears in the denominator of $\not{s}_1$. Only one spin projection operator is needed, because as we have seen in the spinor analysis the vector interaction couples the two spins together, so that if the electron is right handed, the positron is automatically left handed. We have two traces, one of which we have seen several times before:

$$\frac{1}{8}\mathrm{Tr}(\not{p}_2\gamma^\mu\not{p}_1\gamma^\nu) + \frac{m}{8}\mathrm{Tr}(\not{p}_2\gamma^\mu\gamma^5\not{s}_1\gamma^\nu). \tag{4.115}$$

This is our first encounter with a trace that involves γ^5, but, alas, it will not be our last! Rearrange terms $\not{p}_2\gamma^\mu\gamma^5\not{s}_1\gamma^\nu = \gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu\gamma^5 p_{2\alpha}s_{1\beta}$. Now refer to Equation (4.48), which we re-write in terms of these indices:

$$\mathrm{Tr}(\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu\gamma^5) = 4\mathrm{i}\epsilon^{\alpha\mu\beta\nu}, \quad \text{where } \epsilon^{0123} = -1. \tag{4.116}$$

$\epsilon^{\alpha\mu\beta\nu}$ vanishes if any two indices are the same, and changes sign if any two indices are interchanged. Reduction formulas for sums of products of four-dimensional ϵ's are given in Chapter 2, and could be used here, but we will simply multiply the terms together. Since the initial state lies along the z axis, the indices α and β can take on values zero and three, and since the indices have to all be different, μ and ν are restricted to one and two. The four vectors are

$$p_1 = (E, 0, 0, E), \quad p_2 = (E, 0, 0, -E), \quad \text{and}$$
$$s_1 = (E/m, 0, 0, E/m). \tag{4.117}$$

Then after a little algebra we have:

$$\frac{m}{8}\mathrm{Tr}(\not{p}_2\gamma^\mu\gamma^5\not{s}_1\gamma^\nu) = iE^2(\delta_{\mu,1}\delta_{\nu,2} - \delta_{\nu,1}\delta_{\mu,2}). \tag{4.118}$$

Equation (4.115) then becomes, after reducing the first trace as we have done before:

$$E^{\mu\nu}_{RL} = E^2(\delta_{\mu,1}\delta_{\nu,1} + \delta_{\mu,2}\delta_{\nu,2} + i(\delta_{\mu,1}\delta_{\nu,2} - \delta_{\mu,2}\delta_{\nu,1})). \tag{4.119}$$

We do the same thing for the final state, calculating the reaction $RL \to RL$, where the second helicities are for the μ^- and μ^+, respectively. The muon tensor then is

$$M^{RL}_{\mu\nu} = \frac{1}{2}\mathrm{Tr}(\not{p}_4\gamma_\mu(\not{p}_3 + m_\mu)(1 + \gamma^5\not{s}_3)\gamma_\nu). \tag{4.120}$$

The four vectors in the final state are more complicated. They are

$$p_3 = (E, E\sin\theta, 0, E\cos\theta), \quad p_4 = (E, -E\sin\theta, 0, -E\cos\theta),$$
$$s_3 = (E/m_\mu, (E\sin\theta)/m_\mu, 0, (E\cos\theta)/m_\mu). \tag{4.121}$$

We are faced with the trace:

$$\frac{m_\mu}{2}\mathrm{Tr}(\gamma_\alpha\gamma_\mu\gamma_\beta\gamma_\nu\gamma^5)p_4^\alpha s_3^\beta = -\frac{4im_\mu}{2}p_4^\alpha s_3^\beta \epsilon^{\alpha\mu\beta\nu}. \tag{4.122}$$

There are six combinations of α and β, but we only need to evaluate two of them, namely, $\alpha = 0,\ \beta = 3$, and $\alpha = 3,\ \beta = 0$, because the imaginary part of the electron tensor, with which the muon tensor is contracted, involves only $\mu = 1,\ \nu = 2$ and $\mu = 2,\ \nu = 1$. The result is

$$-\frac{4im_\mu}{2}p_4^\alpha s_3^\beta \epsilon^{\alpha\mu\beta\nu} = -4iE^2\cos\theta(\delta_{\mu,1}\delta_{\nu,2} - \delta_{\mu,2}\delta_{\nu,1}). \tag{4.123}$$

The contraction of the two tensors gives:

$$E^{\mu\nu}M_{\mu\nu}(RL \to RL) = 4E^4(1 + \cos^2\theta + 2\cos\theta) = 4E^4(1 + \cos\theta)^2. \tag{4.124}$$

The formula for reversing all helicities: $LR \to LR$ is the same, while $LR \to RL = RL \to LR = 4E^4(1 - \cos\theta)^2$, so when you add them all up you get:

$$E^{\mu\nu}M_{\mu\nu} = 4E^4(2(1 + \cos\theta)^2 + 2(1 - \cos\theta)^2) = 16E^4(1 + \cos^2\theta). \tag{4.125}$$

These formulas are given in Peskin and Schroeder, after a slightly different derivation [4]. The helicity conservation amplitudes vanish at $\theta = \pi$, while the flip amplitudes vanish at $\theta = 0$. The polar angle is that of the μ^- in the final state. The μ^- cannot be emitted backwards in an $RL \to RL$ reaction, for that would violate the conservation of angular momentum. The total cross-sections for the various helicity combinations are equal. The cross terms vanish when integrated $0 \leq \theta \leq \pi$.

Using the four vector s^β and its projection operator give helicity amplitudes for any massive spin 1/2 particle that can be described by the Dirac equation. Dirac particles with $m = 0$ have no rest frame, so we have to start over for neutrinos. The appropriate projection operators are $(1 \pm \gamma^5)/2$ introduced in Chapter 2, Section 2.8. We shall return to applications of the massless spin projectors in Chapters 6 and 7. Møller scattering, $e^- + e^- \to e^- + e^-$ elastic scattering, has a very strong dependence on the longitudinal polarizations of both the beam and target electrons, and has been used to measure electron beam polarization. The unpolarized cross-section was obtained in Section 4.9. The polarized cross-section is the subject of Problem 7.1.

We close our discussion of spin effects for now with a justification for the use of the metric tensor $g^{\mu\nu}$ to describe the polarization of the intermediate photon in the photon propagator. The photon polarization and momentum four vectors are

$$\epsilon^\mu = (\epsilon^0, \vec{\epsilon}),\ k^\mu = (\omega, \vec{k}),\ \text{ with } \epsilon \cdot k = 0, \quad \text{and} \quad k \cdot k = 0. \tag{4.126}$$

As we have previously noted, the photon polarization is unchanged by the addition of any complex number c multiplying the four momentum. Thus,

$$\epsilon^\mu \to \epsilon^\mu + c \times k^\mu \text{ leaves } \epsilon \cdot \epsilon \text{ and } \epsilon \cdot k \text{ invariant.} \tag{4.127}$$

The vector interaction couples the photon polarization to some four vector operator, which we shall call M_μ. Since the addition of $c \times k^\mu$ has no effect on the physics, it follows that $k \cdot M = 0$. Now suppose that $\vec{k} = \omega \hat{z}$. Then:

$$k^\mu M_\mu = \omega(M_0 - M_3) = 0, \quad \text{or} \quad M_0 = M_3. \tag{4.128}$$

The sum over physical photon polarization states would be:

$$\begin{aligned} -\sum_{\epsilon^1,\epsilon^2} |M|^2 &= -|M_1|^2 - |M_2|^2 = |M_0|^2 - |M_1|^2 - |M_2|^2 - |M_3|^2 \\ &= M^\mu g_{\mu\nu} M^{\nu *}. \end{aligned} \tag{4.129}$$

The terms $|M_0|^2$ and $-|M_3|^2$ can be added without penalty because they cancel. The metric tensor appears naturally in the photon propagator, and represents the sum over photon polarizations. This is called Ward's identity, and is proven in many texts. This argument follows Appendix B in Nachtmann.

4.11. Compton Scattering $\gamma + e^- \to \gamma + e^-$

Compton scattering is a very important phenomenon in the understanding of quantum electrodynamics. When Arthur Compton did his experiments in 1922 scattering Molybdenum K_α X-rays from electrons in a Carbon target, his interpretation of the change in wave length as a function of scattering angle gave support to the concept of X-rays as massless particles, called photons, that obeyed relativistic kinematics. The geometry is shown in Figure 4.5. The Mo K_α line is a 17.5 keV X-ray, with wave length $\lambda_{K_\alpha} = 0.0709$ nm. Compton measured the wave length of the scattered X-rays by Bragg reflection on a Calcite crystal, and detected the scattered beam in an ionization chamber. His results are shown in Table 4.2 [5].

The quoted uncertainty is an estimate by Compton of the precision of the wave length measurements. The derivation of the Compton formula in introductory classes on modern physics proceeds as

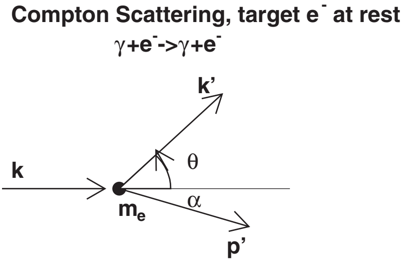


Figure 4.5: Geometry for Compton scattering of X-ray photons off electrons at rest in a Carbon target. The recoil electron in the final state was not observed.

Table 4.2: Compton's wave length data; Mo K_α X-rays.

Scattering angle	Wavelength (nm)	$\Delta\lambda$	Equation (4.131)
0	0.0709 ± 0.0001	0	0
45°	0.0715 ± 0.0001	0.0006	0.0007
90°	0.0731 ± 0.0001	0.0022	0.0024
135°	0.0749 ± 0.0001	0.0040	0.0041

follows:

$$\text{Energy conservation } m_e c^2 + h\nu = \sqrt{(m_e c^2)^2 + (cp')^2} + h\nu',$$
$$\text{Momentum conservation } h\nu/c = (h\nu'/c)\cos\theta + p'\cos\alpha,$$
$$(h\nu'/c)\sin\theta = p'\sin\alpha. \tag{4.130}$$

The photon frequency and wave length are related by $\nu = c/\lambda$. The recoil electron was not observed in the experiment, and can be eliminated from the conservation formulas to obtain:

$$\lambda' - \lambda = \frac{h}{mc}(1 - \cos\theta), \tag{4.131}$$

where $\lambda_c = h/(mc)$ is the Compton wave length. Convert $h/(mc) = hc/(mc^2) = 1240/511,000 = 0.0024$ nm to obtain the numbers in the last column of Table 4.2. The change in wave length $\Delta\lambda$ is independent of the incident photon energy. It is given by the constant Compton wave length, and the cosine of the X-ray scattering angle. Visible light wave lengths are $\lambda \sim 500\,\text{nm}$, so $\Delta\lambda/\lambda \sim 5 \times 10^{-6}$, a very small number. $\Delta\lambda$ can be measured for a 20 keV X-ray, because the K_α line is very sharp, and Bragg reflection gives a precise wave length measurement. The fractional change is much larger for a 200 keV γ ray, with wave length $\lambda = 0.0062\,\text{nm}$, but Bragg reflection is not so convenient to use, because the wave length is much smaller than the crystal spacing. It is a happy coincidence that the wave lengths of X-rays from most elements are comparable to the plane spacing in crystals. The energy of higher energy γ rays is measured by total absorption in a scintillating crystal, where the resolution is good, but not as precise as Bragg's law. The electron mass sets the scale for the kinematics. For photon energies much below 511 keV the target electron is infinitely heavy, and the photon just bounces off. But if the photon energy is even a few percent of the electron mass, it can transfer measurable energy to the target.

Equation (4.131) in our units reads:

$$\frac{1}{\omega'} - \frac{1}{\omega} = \frac{1}{m}(1 - \cos\theta). \tag{4.132}$$

The Compton formula is kinematics. Compton's experiments preceded Schrödinger's wave mechanics by 3 years, so there was no theoretical basis for calculating the cross-section at that time. The Klein–Nishina formula is that cross-section, describing the electromagnetic interaction between the incident photon and the charged electron. The Klein–Nishina formula was developed in 1928, and published in 1929 [6]. The authors had been interested in applying relativistic quantum mechanics to γe^- scattering, and first used the Klein–Gordon equation (same Klein) to describe the electron. They treated the electromagnetic wave in a semi-classical fashion, using relativistic kinematics, but no photon quantization. A similar approach had been used for radiative transitions in atomic spectra. The Klein–Gordon equation has no spin, so possible coupling of the photon to the electron magnetic moment was not taken into account. When Dirac's relativistic equation for an electron appeared in 1928, Klein and Nishina jumped at the chance to apply it to the Compton effect. The electromagnetic wave was still treated semi-classically, but the charged particle wave function was a Dirac spinor, and Dirac matrices appeared in the electric current density. Their result in our units was:

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{2m^2}\frac{\omega'^2}{\omega^2}\left\{\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta\right\}. \tag{4.133}$$

In contrast to other formulas in this chapter, the Klein–Nishina formula is usually not written in Mandelstam variables, although it can be written in invariant form. Here it is in the 'laboratory' frame, in which the target electron is initially at rest. This is the frame of Compton's experiment. Part of the angular dependence is concealed in the elegant form $\omega'/\omega + \omega/\omega'$, where ω' is given by Compton's formula. The electron mass appears explicitly in the formula as $(\lambda_c/(2\pi))^2 = 1.49 \times 10^{-21}$ cm^2. The units for the cross-section are 1/m^2, the rest of the formula being dimensionless. The fine structure constant is $\alpha = 1/137$. The electron mass cannot be ignored in our calculations.

It took Klein and Nishina about 6 months to obtain their famous result that has stood the test of time, and has been confirmed by the more sophisticated techniques of modern quantum electrodynamics.

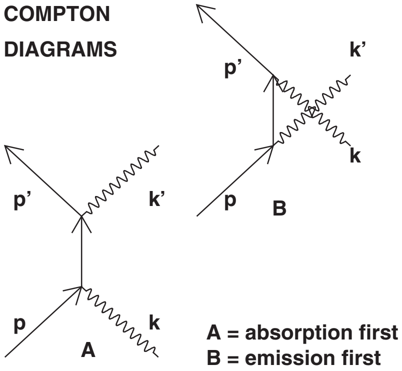


Figure 4.6: Feynman diagrams for Compton scattering. In the lower diagram the incident photon is first absorbed by the target electron, while in the upper one the target electron emits the scattered photon first.

It won't take us 6 months, but even with all of our firepower the cross-section is not trivial to derive. In fact, it is without question the most difficult calculation of this chapter.

The two Feynman diagrams are shown in Figure 4.6. In the lower diagram the incident electron absorbs the incident photon, to form an electron propagator of four momentum $p+k$, that then emits the final state photon and electron $k'+p'$. The upper diagram has the outgoing photon emitted before the incoming photon is absorbed, so that the electron propagator has four momentum $p-k'$. Applying the Feynman rules to the diagrams gives the T matrix element:

$$\begin{aligned} i\langle f|\,T\,|i\rangle &= \bar{u}(p')(-ie\gamma^{\mu})\epsilon^{*}_{\mu}(k')\frac{i(\not{p}+\not{k}+m)}{(p+k)^2-m^2}(-ie\gamma^{\nu})\epsilon_{\nu}(k)u(p) \\ &\quad + \bar{u}(p')(-ie\gamma^{\nu})\epsilon_{\nu}(k)\frac{i(\not{p}-\not{k}'+m)}{(p-k')^2-m^2}(-ie\gamma^{\mu})\epsilon^{*}_{\mu}(k')u(p). \end{aligned} \tag{4.134}$$

The two amplitudes are added because the exchanged photon obeys Bose–Einstein statistics. The various four vectors obey the following relations:

$$p+k=p'+k'; \quad p\cdot p=p'\cdot p'=m^2; \quad k\cdot k=k'\cdot k'=0;$$
$$\epsilon(k)\cdot k=\epsilon^{*}(k')\cdot k'=0. \tag{4.135}$$

The Mandelstam variables are

$$\begin{aligned} s = (p+k)^2 = 2p\cdot k + m^2 = (p'+k')^2; \quad p\cdot k = p'\cdot k'; \\ t = (p-p')^2 = 2m^2 - 2p\cdot p' = (k-k')^2 = -2k\cdot k'; \\ u = (p-k')^2 = (k-p')^2; \\ p\cdot k' = k\cdot p'; \quad s+t+u = 2m^2. \end{aligned} \tag{4.136}$$

The denominators simplify: $(p+k)^2 - m^2 = 2p\cdot k$, and $(p-k')^2 - m^2 = -2p\cdot k'$. The numerators can also be simplified by using some Dirac algebra tricks. In the first term:

$$\begin{aligned} &(\not{p} + \not{k} + m)\gamma^\nu u(p)\epsilon_\nu(k) \\ &\quad = \not{k}\gamma^\nu u(p)\epsilon_\nu(k) + 2p^\nu u(p)\epsilon_\nu(k) - \gamma^\nu(\not{p} - m)u(p)\epsilon_\nu(k) \\ &\quad = (\not{k}\gamma^\nu + 2p^\nu)u(p)\epsilon_\nu(k); \end{aligned} \tag{4.137}$$

the third term being the Dirac equation. The same trick works in the second numerator as well. So, our amplitude becomes:

$$\begin{aligned} i\langle f|T|i\rangle = -ie^2 \frac{\epsilon^*_\mu(k')\epsilon_\nu(k)\bar{u}(p')(\gamma^\mu\not{k}\gamma^\nu + 2\gamma^\mu p^\nu)u(p)}{2p\cdot k} \\ -ie^2 \frac{\epsilon_\nu(k)\epsilon^*_\mu(k')\bar{u}(p')(\gamma^\nu\not{k}'\gamma^\mu - 2\gamma^\nu p^\mu)u(p)}{2p\cdot k'} = A + B. \end{aligned} \tag{4.138}$$

The B term can be transformed into the A term by the replacement $k' \to -k$.

The $\sum |\langle f|T|i\rangle|^2/4$ gives two main terms and a cross term. The factor 1/4 comes from averaging over the initial spin states. Although the photon has spin one, it has only two spin states.

$$\sum_{\text{spins}} \frac{|\langle f|T|i\rangle|^2}{4} = \frac{1}{4}(|A|^2 + |B|^2 + B^\dagger A + A^\dagger B). \tag{4.139}$$

Look at the A term first:

$$\frac{1}{4}\sum_{\text{spins}}|A|^2$$

$$= \frac{e^4}{16(p\cdot k)^2}\sum_{\text{spins}}\epsilon_\alpha(k')\epsilon^*_\mu(k')\epsilon_\nu(k)\epsilon^*_\beta(k)\bar{u}(p')(\gamma^\mu \not{k}\gamma^\nu + 2\gamma^\mu p^\nu)u(p)$$

$$\times\, \bar{u}(p)(\gamma^\beta \not{k}\gamma^\alpha + 2\gamma^\alpha p^\beta)u(p'). \tag{4.140}$$

Note that in the second line the order of $\gamma^\beta\not{k}\gamma^\alpha$ is reversed as a consequence of Hermitian conjugation. We know that the sums over electron spins can be replaced by the completeness relations. The sums over photon spins can be replaced by metric tensors, according to the Ward identity discussed above. So, we have:

$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}}|A|^2 &= \frac{e^4}{16(p\cdot k)^2}g_{\mu\alpha}g_{\nu\beta}\mathrm{Tr}((\not{p}' + m)(\gamma^\mu\not{k}\gamma^\nu + 2\gamma^\mu p^\nu)\\ &\quad\times(\not{p} + m)(\gamma^\beta\not{k}\gamma^\alpha + 2\gamma^\alpha p^\beta))\\ &= \frac{e^4}{16(p\cdot k)^2}\mathrm{Tr}(\not{p}' + m)(\gamma_\alpha\not{k}\gamma_\beta + 2\gamma_\alpha p_\beta)(\not{p} + m)\\ &\quad\times(\gamma^\beta\not{k}\gamma^\alpha + 2\gamma^\alpha p^\beta))\\ &= \frac{e^4}{16(p\cdot k)^2}\mathrm{Tr}((\not{p}'\gamma_\alpha\not{k}\gamma_\beta + 2\not{p}'\gamma_\alpha p_\beta + m\gamma_\alpha\not{k}\gamma_\beta + 2m\gamma_\alpha p_\beta)\\ &\quad\times(\not{p}\gamma^\beta\not{k}\gamma^\alpha + 2\not{p}\gamma^\alpha p^\beta + m\gamma^\beta\not{k}\gamma^\alpha + 2m\gamma^\alpha p^\beta)).\end{aligned} \tag{4.141}$$

There are 16 terms. Half of them are linear in the electron mass and have an odd number of γ matrices, and therefore vanish. So we need to look at terms without the mass, and proportional to m^2. The mass independent terms are

$$\begin{aligned}&\mathrm{Tr}(\not{p}'\gamma_\alpha\not{k}\gamma_\beta\not{p}\gamma^\beta\not{k}\gamma^\alpha\\ &\qquad + 2\not{p}'\gamma_\alpha\not{k}\gamma_\beta\not{p}\gamma^\alpha p^\beta + 2\not{p}'\gamma_\alpha p_\beta\not{p}\gamma^\beta\not{k}\gamma^\alpha + 4\not{p}'\gamma_\alpha p_\beta\not{p}\gamma^\alpha p^\beta)\\ &\quad = (32(p'\cdot k)(p\cdot k) - 16m^2(p'\cdot k) - 16m^2(p'\cdot k) - 32m^2(p'\cdot p))\\ &\quad = 32((p'\cdot k)(p\cdot k) - m^2(p'\cdot p + p'\cdot k)).\end{aligned} \tag{4.142}$$

The terms in the second line of the equation correspond to the traces of the four terms in the first line, that is to say that the first trace is $32(p'\cdot k)(p\cdot k)$. In evaluating the traces we have used $\not{p}\not{p} = p\cdot p = m^2$ and $\not{k}\not{k} = 0$, in addition to the various theorems given earlier. No trace is larger than four γ matrices. The m^2 terms are

$$\begin{aligned}
&\frac{e^4m^2}{16(p\cdot k)}\mathrm{Tr}(\gamma_\alpha\not{k}\gamma_\beta\gamma^\beta\not{k}\gamma^\alpha + 2\gamma_\alpha\not{k}\gamma_\beta\gamma^\alpha p^\beta + 2\gamma_\alpha p_\beta\gamma^\beta\not{k}\gamma^\alpha \\
&\qquad + 4\gamma_\alpha p_\beta\gamma^\alpha p^\beta) \\
&\quad = \frac{e^4m^2}{16(p\cdot k)}(0 + 32(k\cdot p) + 32(k\cdot p) + 64m^2). \qquad (4.143)
\end{aligned}$$

So, finally we have the answer:

$$\begin{aligned}
\frac{1}{4}\sum_{\text{spins}}|A|^2 = \frac{2e^4}{(p\cdot k)^2}&((p'\cdot k)(p\cdot k) + 2m^2(k\cdot p) \\
&- m^2((p'\cdot p) + (p'\cdot k)) + 2m^4). \qquad (4.144)
\end{aligned}$$

The $|B|^2$ term can be obtained directly by substituting $k \to -k'$:

$$\begin{aligned}
\frac{1}{4}\sum_{\text{spins}}|B|^2 = \frac{2e^4}{(p\cdot k')^2}&((p'\cdot k')(p\cdot k') \\
&- 2m^2(k'\cdot p) + m^2((p'\cdot k') - (p'\cdot p)) + 2m^4). \qquad (4.145)
\end{aligned}$$

Since $k\cdot p = k'\cdot p'$, and $k\cdot p' = k'\cdot p$, we may write the sum in a form that resembles Klein and Nishina:

$$\begin{aligned}
\frac{1}{4}\sum_{\text{spins}}(|A|^2 + |B|^2) = 2e^4\Bigg\{&\frac{p'\cdot k}{p\cdot k} + \frac{p\cdot k}{p\cdot k'} + m^2\left(\frac{1}{p\cdot k} - \frac{1}{p\cdot k'}\right) \\
&+ m^4\left(\frac{1}{(p\cdot k)^2} + \frac{1}{(p\cdot k')^2}\right)\Bigg\}. \qquad (4.146)
\end{aligned}$$

The target electron is at rest in the Klein–Nishina formula, and that means that the initial four momentum $p = (m, 0, 0, 0)$. So, $p\cdot k = m\omega$, and $p\cdot k' = p'\cdot k = m\omega'$, and the first two terms are ω'/ω and ω/ω'. This should give us courage to soldier on.

We still have the cross terms. I said that this was not easy. The trace of an even number of γ matrices is invariant under the reversal of the order of the operators (not true for matrices in general), so the two traces $A^\dagger B$ and $B^\dagger A$ are the same. Twice the spin averaged cross term $A^\dagger B$ is

$$\begin{aligned}\frac{1}{2}\sum_{\text{spins}} A^\dagger B &= \frac{e^4}{8(p\cdot k)(p\cdot k')}\sum_{\text{spins}} \epsilon_\alpha(k)\epsilon^*_\nu(k)\epsilon_\mu(k')\epsilon^*_\beta(k')\bar{u}(p)\\ &\quad\times(\gamma^\nu \not{k}\gamma^\mu + 2\gamma^\mu p^\nu)u(p')\bar{u}(p')(\gamma^\alpha \not{k}'\gamma^\beta - 2\gamma^\alpha p^\beta)u(p)\\ &= \frac{e^4}{8(p\cdot k)(p\cdot k')}\mathrm{Tr}\{(\not{p}+m)(\gamma^\nu\not{k}\gamma^\mu + 2\gamma^\mu p^\nu)(\not{p}'+m)\\ &\quad\times(\gamma_\nu\not{k}'\gamma_\mu - 2\gamma_\nu p_\mu)\}. \qquad (4.147)\end{aligned}$$

Again there are 16 terms, and eight of them (linear in the electron mass) vanish, leaving four terms that are mass independent, and four that are proportional to m^2. They are

$$\begin{aligned}&= \frac{e^4}{8(p\cdot k)(p\cdot k')}(\mathrm{Tr}\,\{\not{p}\gamma^\nu\not{k}\gamma^\mu\not{p}'\gamma_\nu\not{k}'\gamma_\mu + 2\not{p}\gamma^\mu p^\nu\not{p}'\gamma_\nu\not{k}'\gamma_\mu\\ &\quad - 2\not{p}\gamma^\nu\not{k}\gamma^\mu\not{p}'\gamma_\nu p_\mu - 4\not{p}\gamma^\mu p^\nu\not{p}'\gamma_\nu p_\mu\}\\ &\quad + m^2\mathrm{Tr}\,\{\gamma^\nu\not{k}\gamma^\mu\gamma_\nu\not{k}'\gamma_\mu - 2\gamma^\nu\not{k}\gamma^\mu\gamma_\nu p_\mu + 2\gamma^\mu p^\nu\gamma_\nu\not{k}'\gamma_\mu\\ &\quad - 4\gamma^\mu p^\nu\gamma_\nu p_\mu\}). \qquad (4.148)\end{aligned}$$

Evaluating the traces in order gives:

$$\begin{aligned}&= \frac{e^4}{8p\cdot kp\cdot k'}\left\{-32k\cdot k'p\cdot p' - 16(2p\cdot k'p\cdot p' - m^2k'\cdot p')\right.\\ &\quad\left. + 16(2p\cdot kp\cdot p' - m^2p'\cdot k) - 16m^2p\cdot p'\right\}\\ &\quad + \frac{e^4m^2}{8p\cdot kp\cdot k'}\left\{16k\cdot k' - 32k\cdot p + 32p\cdot k' - 16m^2\right\}. \qquad (4.149)\end{aligned}$$

Using the conservation laws, the first four terms reduce to $-16m^4$, as several terms cancel. The second four terms become $-16(k\cdot k' + m^2)$.

Assembling all of the parts gives:

$$\frac{1}{4}\sum_{\text{spins}}(A^\dagger + B^\dagger)(A+B) = 2e^4\left\{\frac{k'\cdot p}{k\cdot p} + \frac{k\cdot p}{k'\cdot p} + 2m^2\left(\frac{1}{k\cdot p} - \frac{1}{k'\cdot p}\right)\right.$$
$$\left. + m^4\left(\frac{1}{k\cdot p} - \frac{1}{k'\cdot p}\right)^2\right\}. \tag{4.150}$$

It is amazing that using Compton's formula $1/\omega' - 1/\omega = (1 - \cos\theta)/m$ in this expression plus $k\cdot p = m\omega$ and $k'\cdot p = m\omega'$ gives most of the Klein–Nishina formula:

$$\frac{1}{4}\sum_{\text{spins}}(A^\dagger + B^\dagger)(A+B) = 2e^4\left\{\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta\right\}. \tag{4.151}$$

All that remains is the flux and phase space factors to get the cross-section. Refer to Equation (4.10). The flux term is $1/(4m\omega)$. Since the recoil electron is not observed, integrate over $\delta(\vec{k} - \vec{k}' - \vec{p}')d^3p'$. This leaves us with phase space factors:

$$\frac{(2\pi)^4}{(2\pi)^6}\frac{d^3k'}{2\omega' 2E' 4m\omega}\delta(m + \omega - \omega' - E');$$
$$m + \omega - \omega' - E' = m + \omega - \omega' - \sqrt{m^2 + \omega^2 + \omega'^2 - 2\omega\omega'\cos\theta}$$
$$= f(\omega'). \tag{4.152}$$

Spherical coordinates in k' space gives the integral over ω':

$$\frac{1}{(2\pi)^2}\frac{\omega' d\omega' d\Omega}{16mE'\omega}\delta(f(\omega')); \quad df/d\omega' = -1 - \frac{\omega' - \omega\cos\theta}{E'}. \tag{4.153}$$

Now use Compton's formula:

$$\omega\cos\theta = \omega - \frac{m(\omega - \omega')}{\omega'}; \quad -E' - \omega' + \omega\cos\theta = -\frac{m\omega}{\omega'}; \ \frac{d\omega'}{df} = \frac{E'\omega'}{m\omega}. \tag{4.154}$$

We have changed the sign; the δ function is an even function. The Compton cross-section with the phase space and flux terms then becomes:

$$\frac{d\sigma(\gamma + e^- \to \gamma + e^-)}{d\Omega} = \frac{2\omega'^2 e^4}{(2\pi)^2 16 m^2 \omega^2}\left\{\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta\right\}. \tag{4.155}$$

Substituting $e^4 = 16\pi^2\alpha^2$ gives Equation (4.133). The classical electron radius, written in cgs units as $r_0 = e^2/(mc^2) = 2.8 \times 10^{-13}\,\mathrm{cm}$ is in our units $r_0 = \alpha/m$. Many electromagnetic cross-sections can be expressed in terms of r_0^2. For the Klein–Nishina formula:

$$\frac{d\sigma}{d\Omega} = \frac{r_0^2}{2}\frac{\omega'^2}{\omega^2}\left\{\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta\right\}. \tag{4.156}$$

The total cross-section is well behaved. There is no Coulomb pole, as the photon is not charged. In the low energy limit the Compton wave length is negligible, so $\omega' = \omega$. This gives the Thomson cross-section for the scattering of classical radiation by a free electron.

$$\sigma_T(\omega' = \omega) = \pi r_0^2 \int_{-1}^{+1} dx(1 + x^2) = \frac{8\pi r_0^2}{3}, \tag{4.157}$$

where $x = \cos\theta$. To evaluate the cross-section in the center of mass in the high energy limit where $s \gg m^2$, return to the invariant form Equation (4.150), substitute the Mandelstam variables for the four vector dot products, and use the handy formula Equation (4.69) to evaluate the phase space and flux in the center of mass in the high energy limit:

$$\frac{d\sigma(s \gg m^2)}{d\Omega} = -\frac{\alpha^2}{2s}\left\{\frac{s}{u} + \frac{u}{s}\right\}. \tag{4.158}$$

The variable u is negative. At high energy in the center of mass, where all energies are the same, $u \to -s(1 + \cos\theta)/2$, but the singularity at $\theta = \pi$ is not really there. The electron mass comes to the rescue. In this limit the Compton cross-section resembles other formulas that we have derived in this chapter.

The Klein–Nishina formula is for unpolarized photons and unpolarized electrons. There are however spin effects that can be calculated using the techniques discussed in Section 4.10 for the target electron, and keeping an explicit form for the incident photon polarization. If the recoil electron spin and the scattered photon polarization are not observed, those variables would still be summed over. If RL signifies a right circularly polarized photon incident on an electron at rest with spin pointing opposite to the photon momentum, then a little work will show that $\sigma(RL)$ and $\sigma(RR)$ are different. This difference has been used to measure photon circular polarization by Compton absorption in magnetized iron. Iron has 26 electrons, but only two are polarized by the external magnetic field, and that dilutes the effect to a few percent, but the technique works. The spin dependent term is not too complicated. We quote from an experimental paper on the effect [7]:

$$\frac{d\sigma}{d\Omega} = \frac{r_0^2\omega'^2}{2\omega^2}\left\{\left(\frac{\omega}{\omega'} + \frac{\omega'}{\omega} - \sin^2\theta\right) \pm \left(\left(\frac{\omega}{\omega'} - \frac{\omega'}{\omega}\right)\cos\theta\cos\psi \right.\right.$$
$$\left.\left. + \left(1 - \frac{\omega'}{\omega}\right)\sin\theta\sin\psi\sin\phi\right)\right\}. \tag{4.159}$$

The first part is the unpolarized cross-section. The upper sign is for a photon with positive helicity, called ‘right circularly polarized’ in particle physics. The opposite convention is commonly used in optics, where ‘right circular polarized’ means that the electric vector is rotating clockwise as you look at the beam coming towards you, and we would call that negative helicity. The scattering is in the (x, z) plane, with the incident photon along the +z axis. The free electron is at rest, with its spin oriented at polar angle ψ and azimuthal angle ϕ. Thus, electron spin along the incident photon direction is $\psi = 0$. The formula states that $\sigma(j = 3/2) > \sigma(j = 1/2)$, which is correct, and a way to avoid confusion regarding photon spin orientation.

The total cross-section $\sigma = \sigma_0 \pm \sigma_1$ can be calculated by integrating the two parts of Equation (4.159) over $0 \leq \theta \leq \pi$. The integrals are all elementary, but there are a lot of terms, and it is easy to make a mistake. Introduce the dimensionless energy variable $k = \omega/m$, so that k is not the wave number, but the photon energy divided by the electron mass. The scattered energy in terms of this variable is: $\omega' = \omega/(1 + k(1 - x))$, where $x = \cos\theta$. Then the integrals are

$$\sigma_0 = \frac{\pi r_0^2}{k}\left\{\frac{k^2 - 2k - 2}{k^2}\ln(1 + 2k) + \frac{1}{2} + \frac{4}{k} - \frac{1}{2(1 + 2k)^2}\right\}, \tag{4.160}$$

and

$$\sigma_1 = \frac{\pi r_0^2}{k}\left\{\frac{1 + k}{k}\ln(1 + 2k) - \frac{2(1 + 4k + 5k^2)}{(1 + 2k)^2}\right\}. \tag{4.161}$$

$\pi r_0^2 = 244$ mb. The experiment of Grant and Page referred to above used 2.62 MeV γ rays from a radioactive source to study the transmission through polarized iron. Substituting $k = 5.12$ into these equations gives the cross-sections $\sigma_0 = 122$ mb, and $\sigma_1 = 21$ mb. So, for 100% right circularly polarized photons, the cross-section for electrons polarized along the incoming photon direction would be $\sigma_0 + \sigma_1 = 142$ mb, and for the opposite electron polarization $\sigma_0 - \sigma_1 = 101$ mb, for a 17% effect. This would be diluted by the $2/26 = 0.077$ number of iron electron spins that cooperate with the magnetic field, the degree of photon polarization, and other factors. The cross-section ratio has to be re-evaluated for each γ energy. Nevertheless, it has proven a practical method for measuring photon circular polarization.

Another interesting application of Compton scattering is the back scattering of laser photons by a 50 GeV electron beam. You shine visible light that could be polarized directly at an electron beam that could also be polarized, and get polarized γ rays coming back at you! The kinematics is left as an exercise in the Problems. The technique has been applied to measurements of electron longitudinal polarization in the SLAC Linear Collider, discussed in Section 7.7. Z neutral weak bosons with a mass of 91 GeV were produced by colliding polarized electrons with unpolarized positrons, and the electron polarization was measured using Compton scattering by a polarized laser beam.

Compton scattering is treated in many texts. The derivations presented here were all worked out by the author, but follow the approach given by Peskin and Schroeder. Bjorken and Drell do not sum over photon polarizations, and, without the handy metric tensors that simplify the expression, confront traces of eight γ matrices. Their result contains a dot product of incident and scattered photon polarizations.

We will use the trace techniques introduced in this chapter for the rest of the book, so the reader should get comfortable. Typically there are several cancellations in any trace calculation. The author has often wondered whether some clever person will figure out a method that avoids writing down the cancelling terms. Such an improvement would decrease the probability of error.

Quantum electrodynamics is a very successful theory. Only the basic calculations have been presented in this book. Radiative corrections are difficult to calculate, but have been confirmed as correct by experiment. The Lamb shift splitting of the $2P^{1/2}$ and $2S^{1/2}$ levels of hydrogen mentioned earlier can be calculated as a result of several radiative corrections to the $n = 2$ bound state. The dominant effect is the emission and reabsorption of virtual photons by the electron in its orbit. Another fundamental calculation is the electron anomalous magnetic moment. The Dirac equation as we have seen predicts the g-factor equal to two, defined by the formula for the magnetic moment in terms of the spin as

$$\vec{\mu} = \frac{ge}{2m}\vec{S}. \tag{4.162}$$

The Larmor frequency for the precession of a magnetic dipole in a perpendicular constant magnetic field is

$$\omega_L = \frac{geB}{2m\gamma}, \tag{4.163}$$

where $\gamma = E/m$, the Lorentz factor. If $g = 2$, this frequency is the same as the cyclotron frequency for the rotation of the momentum vector, so if $\vec{\sigma} \cdot \hat{p} = +1$ upon injection into the field, the two vectors will remain aligned as the electron orbits in a circle in the magnetic field. If $g \neq 2$, then the frequencies are different, and the correct relativistic formula, quoted in Chapter 2, is

$$\omega_L - \omega_c = \omega_c \gamma \frac{g-2}{2}. \tag{4.164}$$

The g factor anomaly is defined by $a = (g - 2)/2$. The anomaly can be calculated from the second-order correction to the electron vertex function. This calculation is done in Chapter 6 of Peskin and Schroeder. Many radiative corrections are plagued by divergences either at the highest energies — ultraviolet divergences, or as $\omega \to 0$ — infrared divergences, and sometimes both; but for the anomalous magnetic moment the result is finite:

$$a_e = \frac{\alpha}{2\pi} = 0.0011614. \tag{4.165}$$

This small quantity is multiplied by γ, so the change in angle between the spin and momentum vectors is

$$\Delta\phi = \theta_c \times \gamma \times 0.0011614. \tag{4.166}$$

For a 50 GeV electron beam $\gamma \sim 10^5$, so θ_c is multiplied by roughly 100, and to rotate the spin by 90° from longitudinal to transverse, you need bend the beam by only about 1°. The muon is a lepton obeying the Dirac equation, and its g factor should be very close to that of the electron. The experimental number from the Brookhaven muon $g - 2$ experiment [8] is

$$a_{\mu^+} = 0.0011659191 \pm (59 \times 10^{-10}). \tag{4.167}$$

These anomalies are very useful in experimental physics.

Radiative corrections are treated in most advanced texts. There are many ways to handle the divergences that follow from loop integration. One method is to introduce a fictitious photon mass, or a low energy cut-off below which real photons are not detected. Loop integrals are typically of the form $\int k^2 dk/k^3 = \log(k)$, that diverge both at $k = 0$ and $k = \infty$. But if the integral were $\int k dk/k^3$, a two-dimensional rather than three-dimensional integral, then the ultraviolet divergence would be eliminated. This idea led to perhaps the most elegant way to handle the divergences invented by 't Hooft and Veltman, and called *dimensional regularization.* In this scheme, the calculations are carried out in d dimensions (one time dimension and $d - 1$ space dimensions). If we set $d = 4 - \epsilon$, then we obtain factors of $1/\epsilon$ after integration, that are canceled by factors of ϵ in the final answer. It is amazing that one can do Dirac algebra in d dimensions, where at least one of the space indices is not an integer. In any event, such sophisticated calculations are beyond the scope of this book.

4.12. Problems

4.1 Show the equivalence of various expressions for the flux in the cross-section formula Equations (4.11) and (4.12).

4.2 Derive Equation (4.25) starting with (4.19)–(4.21). Explain how the machinery works.

4.3 Derive the formula for the T matrix element, Equation (4.28) from Equation (4.25).

4.4 Derive Equation (4.33), showing how the sum over spins becomes a trace.

4.5 Prove the identity $\gamma^\mu \not{a}\not{b}\gamma_\mu = 4a \cdot b$. From the definition of γ^5 justify the trace Equation (4.48), by choosing an order for μ, ν, ρ, σ and evaluating the product.

4.6 Write down the muon tensor $M_{\mu\nu}$ and contract it with the electron tensor Equation (4.49), to obtain Equation (4.50).

4.7 Verify the high energy limit formulas in the center of mass, Equations (4.52)–(4.54).

4.8 Work out the explicit spinor matrix elements Equations (4.50)–(4.61), and obtain the electron tensor

$$E^{\mu,\nu} = \frac{1}{4}\sum_{\text{spins}} \bar{v}^{s_2}\gamma^\mu u^{s_1}\bar{u}^{s_1}\gamma^\nu v^{s_2} = 2E^2(\delta_{\mu,1}\delta_{\nu,1} + \delta_{\mu,2}\delta_{\nu,2}).$$

Obtain the same tensor from Equation (4.49) for $p_1^0 = p_2^0 = E$, and $p_1^3 = -p_2^3 = E$.

4.9 Using the same techniques derive the muon tensor Equation (4.66), and the contraction of the two $E^{\mu\nu}M_{\mu\nu}$ Equation (4.67).

4.10 Obtain Equation (4.76) for $e^-\mu^-$ scattering from Equation (4.50) for $e^-e^+ \to \mu^-\mu^+$.

4.11 Draw the diagrams for Bhabha scattering, and use the Feynman rules to construct the amplitudes from the diagrams. Verify the trace calculations for the three terms $|A(s)|^2$, $|A(t)|^2$, and $2A^\dagger(t)A(s)$.

4.12 Repeat the steps outlined in Problem 11 for Møller scattering.

4.13 Verify Equation (4.108) for the spin projection operator Λ_+, and show that $\Lambda_+ u^1(p) = u^1(p)$, and $\Lambda_+ u^2(p) = 0$.

4.14 Verify for $\alpha\mu\beta\nu = 0123$ that $\mathrm{Tr}(\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu\gamma^5) = -4i$.

4.15 Derive the electron tensor $E^{\mu\nu}_{RL}$ of Equation (4.119), and the muon tensor $M^{RL}_{\mu\nu}$ of Equation (4.123).

4.16 Contract the two tensors of Problem 4.15 to obtain the angular distribution for polarized leptons.

4.17 Sharpen your pencil, and evaluate the traces found in $(1/4)\sum_{\text{spins}}|A|^2$ for Compton scattering, Equations (4.142) and (4.143).

4.18 Sharpen it again, and evaluate the traces for the cross term $(1/2)\sum_{\text{spins}}A^\dagger B$, Equation (4.149).

4.19 Show that Equation (4.150) becomes the Klein–Nishina Formula Equation (4.151).

4.20 Perform the integrals for the unpolarized and polarized Compton cross-sections to obtain Equations (4.160) and (4.161).

4.21 Compton back scattering of laser light. Shine a $\lambda = 500\,\mathrm{nm}$ laser (blue-green light) directly at an incident 50 GeV electron beam. Calculate the energy of the back scattered γ rays in the laboratory. You may work either in the center of mass, with Lorentz transformations in and out, or stay in the laboratory. In either case you have to be careful in your approximations, because the electron mass cannot be ignored. The center of mass is moving in the electron direction. With laser photon momentum k and electron momentum p, the velocity of the center of mass is $v_c = (p-k)/(k+\sqrt{m^2+p^2}) \approx 1 - 2k/p - m^2/(2p^2)$. For the given energies, k and $m^2/(2p)$ are about the same, roughly 2.5 eV. Obtain the energy of the back scattered γ ray in the laboratory:

$$k' = \frac{p}{1 + m^2/(4kp)}.$$

For our numbers, this amounts to $k' = 2p/3$, leaving the scattered electron (still traveling in its original direction) with 1/3 of its original energy. You get 33 GeV γ rays coming back at you! The laser and electron beams have to be very carefully aligned, and the resulting γ ray beam is not very intense, but it can be polarized by polarizing the original laser light, and has been used to study spin dependent effects in photoproduction.

References

[1] MEG Collaboration, *Phys. Rev. Lett.* **110**, 201801 (2013).

[2] O. Nachtmann, *Introduction to Particle Physics* (Springer Verlag, Berlin, 1987).

[3] S. Mandelstam, *Phys. Rev.* **112**, 1344 (1958).

[4] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, New York, 1995).

[5] A. H. Compton, *Phys. Rev.* **21**, 483; **22**, 409 (1923).

[6] O. Klein and Y. Nishina, *Zeitschrift für Physik* **52**, 853 (1929).

[7] S. B. Grant and L. A. Page, *Phys. Rev.* **92**, 970 (1953).

[8] H. N. Brown *et al.*, *Phys. Rev. D* **62**, 091101 (2000).

Chapter 5

Electrodynamics with Protons and Neutrons

5.1. Elastic Electron Scattering in the Laboratory Frame

By the term ‘laboratory frame’ we mean a Lorentz frame where the target particle is initially at rest. This is in contrast to the center of mass frame, where most of the calculations were done in the previous chapter. However, the Klein–Nishina formula for Compton scattering was obtained in the laboratory frame where Compton did his experiments. So, this should not be completely new, and as we did for incoming and outgoing photons, we will use kinematics favoring measurements of the incident and scattered electrons, with the recoil target not observed. The measurement of proton and neutron structure by elastic scattering of electrons or muons off hydrogen and deuterium targets has been extensively studied in fixed target experiments at SLAC in California and JLab in Virginia, among other places, and at the specially built electron–proton collider HERA in Hamburg, Germany. The HERA collider was an asymmetric two-storage ring machine, with 27.5 GeV electrons or positrons scattered by 920 GeV protons, with a center of mass energy of 318 GeV. The motion of the center of mass in the laboratory frame was in the proton direction with $v_c = 0.94$, and the values of $-q^2 = -(k-k')^2$ were typically between 10^3 and $10^4\,\mathrm{GeV}^2$. The HERA experiments were in a different realm from the pioneering fixed target measurements that are the subjects of our initial discussion, but we will return to HERA data later in this chapter.

For an electron beam with energy $E \gg m_e$ incident on a nucleon target, two independent variables may be defined for the final state electron, its energy E', and its scattering angle θ. The energy transfer is called $\nu = E - E'$. We write down the kinematics in the laboratory frame where the initial target four momentum is just the proton mass $p = (M, 0, 0, 0)$. The electron four momenta are $k \to k'$, and the target four momenta are $p \to p'$. For elastic scattering, $p^2 = p'^2 = M^2$. For inelastic scattering, the invariant mass of the recoil final state $W^2 > M^2$.

$$q^2 = (k-k')^2 = -2k \cdot k'; \quad k + p = k' + p'; \quad q + p = p';$$
$$q^2 + 2M(E - E') + M^2 = W^2. \tag{5.1}$$

For elastic scattering, M^2 and W^2 cancel in the last equation, leaving

$$-\frac{q^2}{2M\nu} = 1. \tag{5.2}$$

The familiar four-dimensional delta function integral for elastic scattering is

$$\int \frac{d^3p'}{2p'_0}\delta^4(p + q - p') = \frac{\delta(M + \nu - p'_0)}{2p'_0} = \frac{\delta(\nu + q^2/2M)}{2M}. \tag{5.3}$$

For inelastic scattering, the parameter x is defined as the ratio

$$x = -\frac{q^2}{2M\nu}; \quad 0 \le x \le 1. \tag{5.4}$$

The mass in the denominator is the initial state target mass (usually the proton mass), not the recoil mass given by Equation (5.1). In other words, the four momentum transfer and the energy transfer for an inelastic process satisfy the inequality $-q^2 < 2M\nu$. This is the definition of Bjorken x, named after J.D. Bjorken, who first formulated this analysis. We shall see in the parton model that x is the momentum fraction carried by the struck quark. The two final electron variables ν and $q^2 = -4EE' \sin^2(\theta/2)$ are independent for inelastic scattering. They can be measured directly in the scattered beam, and used to parametrize the process. The invariant mass of the unobserved recoil is a known quantity in terms of the measured variables of the scattered electron beam, even if the recoil debris is not observed. As given above: $W^2 = (p+q)^2 = M^2 + 2M\nu + q^2$.

As orientation regarding the problem of proton structure, we return to $e^-\mu^-$ scattering, a cross-section where the muon, a true Dirac particle, is the target. In Chapter 4, we used the shuffling of the kinematic variables to write down the cross-section from the formula for $e^-e^+ \to \mu^-\mu^+$ without doing any work. For our discussion of elastic scattering from the proton, however, we need the two tensors, one the 'electron' tensor $E^{\mu\nu}$ and the other 'muon' tensor $M_{\mu\nu}$. The reason for this is that the electron tensor remains the same, whether the electron is scattered by another Dirac spin 1/2 particle like the muon, or by a compound object like the proton, but the muon tensor has to be modified to allow for our lack of knowledge of the details of the proton structure. So, let us review the derivation of the $e^-\mu^-$ scattering cross-section, beginning with the amplitude from the Feynman rules:

$$\left\langle e^-\mu^- \middle| T \middle| e^-\mu^- \right\rangle = -\frac{e^2}{q^2}(\bar{u}(k')\gamma^\mu u(k))(\bar{u}(p')\gamma_\mu u(p)). \tag{5.5}$$

The spin averaged square is

$$\frac{1}{4}\sum_{\text{spins}} |\langle f|\, T\,|i\rangle|^2 = \overline{|\langle f|\, T\,|i\rangle|^2} = \frac{e^4}{4q^4}E^{\mu\lambda}M_{\mu\lambda}, \tag{5.6}$$

where

$$E^{\mu\lambda} = \mathrm{Tr}(\not{k}'\gamma^\mu \not{k}\gamma^\lambda) = 4(k'^\mu k^\lambda + k'^\lambda k^\mu - (k'\cdot k)g^{\mu\lambda}). \tag{5.7}$$

For the muon tensor, we retain the muon mass term

$$M_{\mu\lambda} = \mathrm{Tr}((\not{p}'+m)\gamma_\mu(\not{p}+m)\gamma_\lambda) = 4(p'_\mu p_\lambda + p'_\lambda p_\mu + (m^2 - (p'\cdot p))g_{\mu\lambda}). \tag{5.8}$$

These are both symmetric tensors $E^{\mu\lambda} = E^{\lambda\mu}$, and because of current conservation (recall the argument for using the metric tensor as the photon spin sum),

$$q_\mu E^{\mu\lambda} = q_\lambda E^{\mu\lambda} = q^\mu M_{\mu\lambda} = q^\lambda M_{\mu\lambda} = 0, \tag{5.9}$$

as can be demonstrated by direct substitution; $q^\mu = k^\mu - k'^\mu$. The contraction of these two tensors gives Equation (4.76) for $e^-\mu^-$ scattering, rewritten in terms of $k + p \to k' + p'$ as

$$\frac{1}{4}\sum_{\text{spins}} |\langle f|\,T\,|i\rangle|^2 = \frac{8e^4}{q^4}\left\{(k\cdot p')(k'\cdot p) + (k\cdot p)(k'\cdot p') - m^2(k\cdot k')\right\}, \tag{5.10}$$

where m is the muon mass. The formula is simplified by using $k'\cdot p = mE' = k\cdot p'$, and $k\cdot p = mE = k'\cdot p'$. The form that is useful for generalization to a composite target is

$$\overline{|\langle f|\,T\,|i\rangle|^2} = \frac{8e^4}{q^4}2m^2EE'\left\{\cos^2(\theta/2) - \frac{q^2}{2m^2}\sin^2(\theta/2)\right\}. \tag{5.11}$$

The bar denotes the average over initial state spins and the sum over final state spins. To calculate the cross-section in the laboratory frame, we rewrite Equation (4.10) in the present notation:

$$d\sigma = (2\pi)^4\delta^4(p + k - p' - k')\frac{d^3k'}{(2\pi)^3 2E'}\frac{d^3p'}{(2\pi)^3 2p_0'}\overline{|\langle f|\,T\,|i\rangle|^2}\frac{1}{4(mE)}, \tag{5.12}$$

where the last factor is the flux. Integrate over the unobserved recoil momentum, leaving an energy δ function. Then integrate $\delta(m + E - E' - p_0')dE'$ to give

$$\frac{d\sigma}{d\Omega} = \frac{1}{64(\pi)^2m^2}\frac{E'^2}{E^2}\overline{|\langle f|\,T\,|i\rangle|^2}. \tag{5.13}$$

Evaluating the argument of the energy δ function gives a factor $(m + E(1 - \cos\theta))$ in the denominator that has been eliminated by the kinematic relation:

$$\frac{E'}{E} = \frac{1}{1 + \frac{2E\sin^2(\theta/2)}{m}}. \tag{5.14}$$

Combining Equations (5.11) and (5.13) gives the cross-section for $e^-\mu^-$ scattering in the laboratory frame, in which the target muon is at rest:

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2\sin^4(\theta/2)}\frac{E'}{E}\left\{\cos^2(\theta/2) - \frac{q^2}{2m^2}\sin^2(\theta/2)\right\}. \tag{5.15}$$

An alternate way of writing this formula is to factor out $\cos^2(\theta/2)$ as follows:

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2\sin^4(\theta/2)}\frac{E'}{E}\cos^2(\theta/2)\left\{1 - \frac{q^2}{2m^2}\tan^2(\theta/2)\right\}. \tag{5.16}$$

Then the part in front of the curly brackets is the Mott cross-section for electron elastic scattering from an infinitely heavy point charge in the relativistic limit $E \gg m_e$. For lower electron velocity $v < 1$, the correct Mott formula is

$$\sigma_M = \frac{\alpha^2}{4(kv)^2\sin^4(\theta/2)}(1 - v^2\sin^2(\theta/2)). \tag{5.17}$$

Here, k is the electron momentum. This form as $v \to 0$ becomes Rutherford's non-relativistic formula for scattering α particles off a heavy nucleus of atomic number Z, after multiplying α^2 by $4\mathrm{Z}^2$. In the non-relativistic limit, $(kv)^2 \to m^2v^4 = 4E_{kin}^2$. For Mott scattering, $E' = E$. That the cross-section vanishes at $\theta = \pi$ as $v \to 1$ is a consequence of helicity conservation.

5.2. Electron–Proton Elastic Scattering

How should we modify $M_{\mu\lambda}$ for a proton/neutron target to describe electron–proton or electron–neutron elastic scattering? First, write $M_{\mu\lambda}$ in terms of the spinors:

$$M_{\mu\lambda} = \sum_{\text{spins}}(\bar{u}(p')\gamma_\mu u(p)) \times (\bar{u}(p)\gamma_\lambda u(p')). \tag{5.18}$$

Now appeal to the Gordon decomposition of the vector current from Section 2.9 of Chapter 2. Equation (2.201) in the present notation reads as

$$e\bar{u}(p')\gamma^\mu u(p) = \frac{e}{2M}\left\{\bar{u}(p')((p'+p)^\mu + i\sigma^{\mu\nu}q_\nu)u(p)\right\}, \tag{5.19}$$

where

$$q_\nu = (k - k')_\nu = (p' - p)_\nu; \quad \text{and} \quad \sigma^{\mu\nu} = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu). \quad (5.20)$$

Here and in the following formulas, M is the proton mass. The proof of the Gordon decomposition given in Chapter 2 is based on the Dirac equations $(\not{p} - M)u(p) = 0$ and $\bar{u}(p')(\not{p}' - M) = 0$. The first term in Equation (5.19) is the transition charge density $e\bar{u}(p')u(p)$ multiplied by the sum of the two four momenta, while the second term is the magnetic moment density multiplied by the difference in the two four momenta. Substituting Equation (5.19) into Equation (5.18) for each of the two spinor products, and taking the trace gives Equation (5.8) for the tensor $M_{\mu\lambda}$. The Gordon decomposition gives us a clue regarding the modification of the pure Dirac current for a composite spin 1/2 particle like the proton or neutron. One way is to introduce two form factors that are functions of q^2:

$$\gamma^\mu \to F_1(q^2)\gamma^\mu + \frac{\kappa}{2M}F_2(q^2)(i\sigma^{\mu\nu}q_\nu), \quad (5.21)$$

where $\kappa = \frac{g}{2} - 1 = 1.79$ is the anomalous part of the proton magnetic moment, and $\kappa = -1.91$ for the neutron. The neutron magnetic moment is all anomalous. For the proton, $F_1(0) = 1$, and for the neutron which has no net charge, $F_1(0) = 0$. For both proton and neutron $F_2(0) = 1$. As $|q^2|$ increases from zero, the form factors are expected to decrease, indicating that the extended charge and magnetic moment distributions are less effective in scattering electrons than a point Dirac particle. Comparing Equation (5.21) to Equation (5.19) shows us that the form factor $F_1(q^2)$ takes care of the transition charge density and the magnetic density that depends on the pure Dirac moment $g = 2$. The second term proportional to $F_2(q^2)$ depends only on the anomalous magnetic moment.

While it is possible to handle the traces involving the product $\sigma^{\mu\nu}$, it is prudent to use the Gordon decomposition to eliminate that operator. To do this, add and subtract a term involving κF_2 as follows:

$$\gamma^\mu \to F_1\gamma^\mu + \frac{\kappa F_2}{2M}iq_\nu\sigma^{\mu\nu} + \frac{\kappa F_2}{2M}(p + p')^\mu - \frac{\kappa F_2}{2M}(p + p')^\mu. \quad (5.22)$$

The second and third terms can be absorbed on the first term to give the baryon current:

$$\bar{u}(p')((F_1+\kappa F_2)\gamma^\mu - \frac{\kappa F_2}{2M}(p+p')^\mu)u(p). \tag{5.23}$$

F_1 and F_2 are real functions of q^2, so the baryon tensor has the following traces:

$$\begin{aligned}M^{\mu\lambda} = {} & (F_1+\kappa F_2)^2\mathrm{Tr}((\not{p}'+M)\gamma^\mu(\not{p}+M)\gamma^\lambda) \\ & -(F_1+\kappa F_2)\frac{\kappa F_2}{2M}\mathrm{Tr}((\not{p}'+M)\gamma^\mu(\not{p}+M)(p+p')^\lambda \\ & +(\not{p}'+M)(\not{p}+M)\gamma^\lambda(p+p')^\mu) \\ & +\frac{(\kappa F_2)^2}{(2M)^2}\mathrm{Tr}((\not{p}'+M)(\not{p}+M))(p+p')^\mu(p+p')^\lambda,\end{aligned} \tag{5.24}$$

where M is the proton mass. The electron mass has been neglected in the tensor $E_{\mu\lambda}$. The traces are elementary. Note that in the cross term, only products $M\gamma^\mu\not{p}$ and $M\not{p}'\gamma^\mu$ contribute. Other terms involve an odd number of γ matrices. Also, note that if $F_1 = 1$ and $\kappa = 0$, we retrieve the muon formula. Evaluating the traces gives the hadron tensor:

$$\begin{aligned}M^{\mu\lambda} = {} & 4(F_1+\kappa F_2)^2(p'^\mu p^\lambda + p'^\lambda p^\mu + (M^2 - p\cdot p')g^{\mu\lambda}) \\ & -4(F_1+\kappa F_2)\kappa F_2(p+p')^\lambda(p+p')^\mu \\ & +4\frac{(\kappa F_2)^2}{(2M)^2}(p\cdot p' + M^2)(p+p')^\mu(p+p')^\lambda.\end{aligned} \tag{5.25}$$

Rewrite Equation (5.7) with lower indices to make the following contraction:

$$E_{\mu\lambda} = 4(k_\mu k'_\lambda + k_\lambda k'_\mu - k\cdot k' g_{\mu\lambda}). \tag{5.26}$$

Then we have the final result:

$$\begin{aligned}E_{\mu\lambda}M^{\mu\lambda} = {} & 32(F_1+\kappa F_2)^2(k\cdot p' k'\cdot p + k\cdot p k'\cdot p' - M^2 k\cdot k') \\ & +32\left\{\frac{(\kappa F_2)^2}{(2M)^2}(p\cdot p'+M^2) - \kappa F_2(F_1+\kappa F_2)\right\} \\ & \times\left(k\cdot(p+p')k'\cdot(p+p') - \frac{k\cdot k'}{2}(p+p')^2\right).\end{aligned} \tag{5.27}$$

5.3. The Rosenbluth Formula

The cross-section formula for electron–proton scattering is called the Rosenbluth formula, after Marshall Rosenbluth, who first wrote it down [1] in a form slightly different from the one in current use. Halzen and Martin [2] give the modern version, in terms of the unknown form factors F_1 and F_2:

$$\begin{aligned}\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2\sin^4(\theta/2)}\frac{E'}{E}\Bigg\{&\left(F_1^2 - \frac{\kappa^2 q^2 F_2^2}{4M^2}\right)\cos^2(\theta/2)\\ &- \frac{q^2}{2M^2}(F_1+\kappa F_2)^2\sin^2(\theta/2)\Bigg\}.\end{aligned} \tag{5.28}$$

Equations (5.28) and (5.15) for electron–muon scattering differ only inside the curly brackets. The contraction of the electron and hadron tensors, Equation (5.27), feeds the curly brackets, but the units are not the same, since Equation (5.27) has dimensions E^4. If we factor out $64(M^2EE')$ from Equation (5.27), then the Rosenbluth formula follows:

$$\begin{aligned}M^{\mu\lambda}E_{\mu\lambda} = 64(M^2EE')\Bigg(&\left(F_1^2 - \frac{\kappa^2 q^2}{4M^2}F_2^2\right)\cos^2(\theta/2)\\ &- \frac{q^2}{2M^2}(F_1+\kappa F_2)^2\sin^2(\theta/2)\Bigg).\end{aligned} \tag{5.29}$$

There are so many ways that one can write down the kinematic equations that it is far from obvious that this claim is true. However, it is. Look at the coefficient of F_1^2, for example,

$$\begin{aligned}&32(k\cdot p'k'\cdot p + k\cdot pk'\cdot p' - M^2k\cdot k')\\ &\quad = 64M^2EE'\left(\cos^2(\theta/2) - \frac{q^2}{2M^2}\sin^2(\theta/2)\right).\end{aligned} \tag{5.30}$$

Using $k\cdot p = k'\cdot p' = ME$, $k\cdot p' = ME - k\cdot k'$, and $q^2 = (k-k')^2 = (p'-p)^2 = -2k\cdot k' = -4EE'\sin^2(\theta/2)$, one gets

$$\begin{aligned}&32M^2\left(EE' + E'\frac{q^2}{2M} + E^2 + \frac{q^2}{2}\right)\\ &\quad = 64M^2EE'\left(\cos^2(\theta/2) - \frac{q^2}{2M^2}\sin^2(\theta/2)\right)\end{aligned} \tag{5.31}$$

or

$$1/2 + \frac{q^2}{4ME} + \frac{E}{2E'} + \frac{q^2}{4EE'} = 1 - \left(1 + \frac{q^2}{2M^2}\right)\sin^2(\theta/2). \quad (5.32)$$

Equation (5.32) is true. Other coefficients can be checked in a similar (tedious) manner. Rosenbluth's form is nice, because all of the modifications are inside the curly brackets, and setting $M = m$, $F_1 = 1$, and $\kappa = 0$ at once retrieves the electron–muon formula.

The form factors $F_1(q^2)$ and $F_2(q^2)$ are unknown, and have to be obtained by experiment. Linear combinations of F_1 and F_2 have been introduced to eliminate the cross term in the curly brackets, and measurements of the proton and neutron form factors now refer to these functions. A good review of the experimental situation, including a discussion of various polarization effects — both electron and proton/neutron — is available [3]. The new functions, called G_E and G_M, are defined as follows:

$$G_E = F_1 - \tau\kappa F_2, \quad G_M = F_1 + \kappa F_2, \quad \tau = -\frac{q^2}{4M^2};$$

$$\text{Inverted } F_1 = \frac{G_E + \tau G_M}{1+\tau}, \quad \kappa F_2 = \frac{G_M - G_E}{1+\tau}. \quad (5.33)$$

G_E and G_M are called the electric and magnetic form factor, respectively. At $q^2 = 0, G_E(0) = F_1(0) = 1$, and $G_M(0) = 1 + \kappa = \mu_p$. Substitution into Equation (5.28) gives the Rosenbluth formula in terms of the new functions as follows:

$$\frac{d\sigma}{d\Omega} = \left(\frac{\alpha^2}{4E^2\sin^4(\theta/2)}\right)\frac{E'}{E} \times \left\{\frac{G_E^2 + \tau G_M^2}{1+\tau}\cos^2(\theta/2) + 2\tau G_M^2 \sin^2(\theta/2)\right\}. \quad (5.34)$$

The cross term has canceled out. At low $|q^2| < M^2$, the factor τ is small, so that the measurements are dominated by G_E^2, but at high $|q^2|$, the situation is reversed. If you measure the elastic *e–p* cross-section at a fixed incident electron energy E, and laboratory scattering angle θ, then q^2 is fixed, and you obtain one data point for the curly bracket. A large amount of data covering many beam energies and scattering angles is required to unravel $G_E(q^2)$ and $G_M(q^2)$.

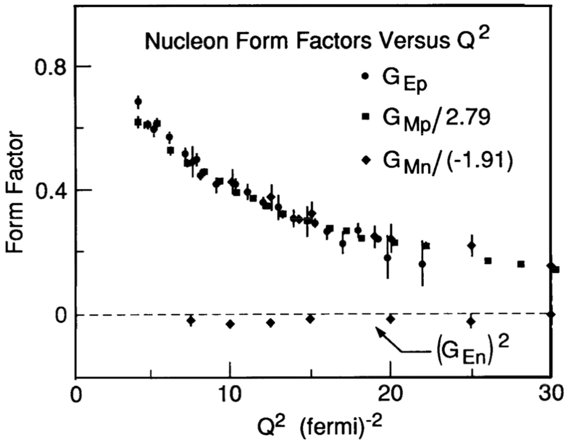


Figure 5.1: Elastic form factors for the proton and neutron extracted from electron scattering from hydrogen and deuterium. This graph, from Stanford data, is adopted from the Nobel Lecture of Richard E. Taylor, 1990.

These functions for the neutron have been measured by quasi-elastic scattering of electrons from deuterium. The magnetic form factors are very similar, that is $G_{Mp}(q^2)/2.79 \sim G_{Mn}(q^2)/-1.91$, and both are equal to $G_{Ep}(q^2)$, while $G_{En}(q^2) \sim 0$. See Figure 5.1.

Both functions are well described for $|q^2| < 20\,\mathrm{GeV}^2$ by what is called a 'dipole' formula:

$$G_D = \frac{1}{(1 - q^2/q_0^2)^2}; \quad q_0 = 0.71\,\mathrm{GeV}. \tag{5.35}$$

The pion Compton wave length $\frac{\lambda}{2\pi} = 2.8 \times 10^{-13}$ cm, so q_0 corresponds to a characteristic distance about five times shorter. If there were no proton recoil, that is if $E = E'$, then the Fourier transform of the static charge density is the form factor:

$$F(\vec{q}) = \int \rho(\vec{x}) e^{i\vec{q}\cdot\vec{x}} d^3x; \quad \text{and} \quad \text{if } \rho(r) = \frac{a^2 e^{-ar}}{4\pi r}, \ \int \rho(r) d^3x = 1,$$

$$\text{then } F(q) = \frac{1}{1 + q^2/a^2}. \tag{5.36}$$

Here, $\vec{q}$ is a three vector, not the space-like four vector in Equation (5.35), but the idea is that the formula for G_D is reminiscent

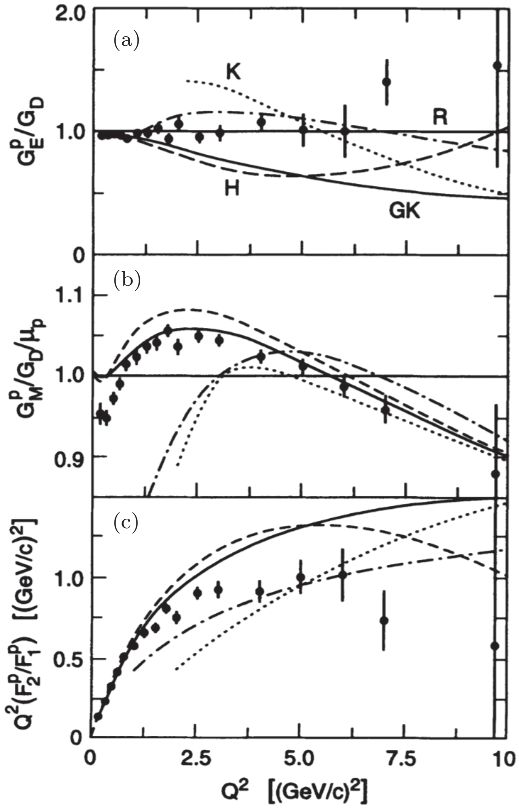


Figure 5.2: All curves are divided by the dipole formula Equation (5.35). The top curve is G_{Ep}/G_D. The middle curve is $(G_{Mp}/2.79)/G_D$, and the bottom curve is $-q^2 \times F_2/F_1$. The x-axis is $Q^2 = -q^2$. The letters on the top curves refer to various theoretical models for deviation from the dipole form. K (dot) and R (dot-dash) are quark models, while GK (solid) and H (dashed) are vector dominance models, where the exchanged photon becomes a hadronic vector meson like ρ^0 that then interacts with the proton. None of the models look good on all three plots. Credit Ref. [4].

of an exponentially damped charge distribution, where the damping distance is about 0.6×10^{-13} cm. Comparison of the experimental data to the dipole Equation (5.35) is shown in Figure 5.2 taken from the measurements of Walker *et al.* [4]. The plots show that, in the $0 < -q^2 < 10\,\mathrm{GeV}^2$ range, the two form factors $G_E(q^2) \sim G_M(q^2)/\mu_p$, to about $\pm 10\%$. This is a remarkable equality, since the charge is a scalar, and the magnetic moment is a vector, and there is no obvious reason why the two constituent sums should be the same. Figure 5.1 shows that this equality of scalar and vector sums does not hold for the neutron, or other neutral baryons that have no charge but large negative magnetic moments like the Λ.

5.4. Electron–Proton Inelastic Scattering

The elastic cross-section at fixed scattering angle θ decreases with increasing $|q^2|$, and becomes quite small for values $-q^2 > 10\,\mathrm{GeV}^2$. For high values of $|q^2|$, the proton breaks up, and is no longer a single baryon, but becomes a recoil baryon plus an additional number of pions. The first excited state is the Δ^+ at 1232 MeV, listed in Table 1.1 of Chapter 1 as a spin 3/2, I-spin 3/2 baryon resonance. The Δ^+ decays with very short lifetime ($\sim 10^{-23}\,\mathrm{s}$) into $p + \pi^0$ and $n + \pi^+$. Other baryon resonances follow, until the cross-section as a function of the recoil invariant mass $W^2 = (p+q)^2$ becomes a smooth curve. See Figure 5.3.

The elastic form factors are evidence of proton structure, but do not give much detailed information. Is it possible to interpret the inelastic scattering at high $|q^2|$ as a point interaction between the exchanged photon and one of the quarks, where the bound quarks are treated as quasi-free Dirac particles, just like the muon but with charge 2/3 or $-1/3$? For this picture to have any chance of working, quark binding must be negligible, and the photon interaction time must be too fast for the struck quark to have any interaction with other quarks or gluons inside the proton. After the photon is absorbed, and the struck quark recoils, the hadronic system must rearrange itself into real colorless mesons and baryons to form the final state, but it has a relatively long time to do this at leisure. The key to this picture is called the infinite momentum frame. It is not really infinite, but it is high enough so that characteristic times for proton binding are lengthened, and the proton size in the direction of motion is contracted (remember flying at the moon in Chapter 2). It is in this frame that the parton model is constructed.

There was a discussion of quark masses in Section 1.16 in Chapter 1. There it was pointed out that there are two different

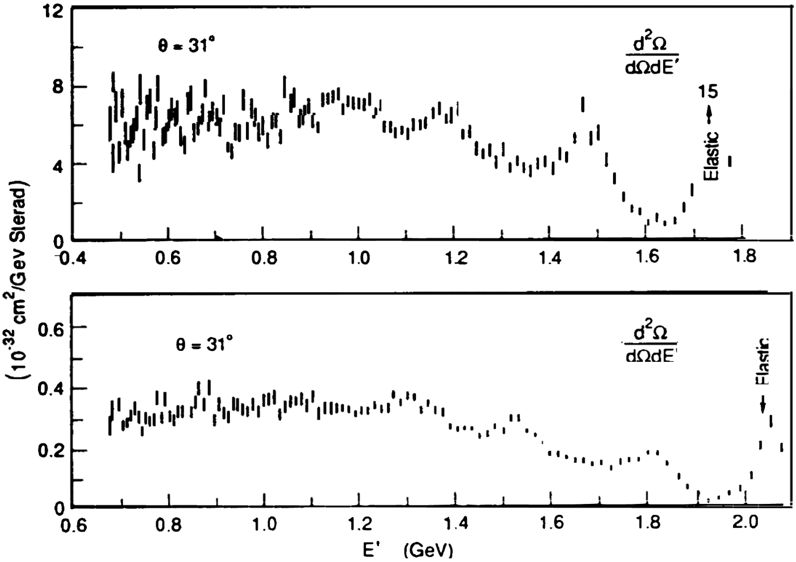


Figure 5.3: Inelastic *e*–*p* scattering in the resonance region. The data are from the Cambridge Electron Accelerator. The top curve is $E_e = 2.4\,\text{GeV}$ and the bottom curve is 3 GeV. The laboratory scattering angle is 31°. The x-axis is the scattered electron energy E', rather than the recoil mass W. Secondary peaks are Δ^+ and higher resonances. The relatively flat and not small continuum at large $\nu = E - E'$ is a clue to the 'deep inelastic' region of the parton model. Credit: the Nobel lecture of Richard E. Taylor, 1990.

mass assignments to the light quarks, the 'constituent quark' masses that are of the order of one-third the proton mass $M/3$, and the 'current quark' masses, that are for the u and d quarks only a few MeV, slightly heavier than the electron. The constituent quarks in the proton and other baryons can be pictured similar to nucleons in, say ^{3}He, which is composed of a bound state of two protons and a neutron. ^{3}He has charge +2e from the protons but has a negative magnetic moment: $\mu_{(^3\text{He})} = -2.13\mu_\text{N}$, about 10% more negative than the neutron! The Pauli principle takes care of the cancellation of the two proton moments, leaving the free neutron to give ^{3}He its spin and magnetic moment. How a negative charge orbital current adds to the neutron moment is not so obvious. However, as we learned in Chapter 1, simple addition of the constituent quark magnetic moments gives a good accounting of the baryon magnetic moments. Constituent quark masses are assumed, each with magnetic moment $\mu_q = e_q/2m_q$, using the color symmetry to determine the total spins for the like flavored quarks, and no orbital angular momentum.

The total binding energy of the ^{3}He nucleus is 7.75 MeV, so the three nucleons viewed from the infinite momentum frame can be considered free. In the helium rest frame

$$M_{(^3\mathrm{He})} = 3M, \tag{5.37}$$

and viewed from the infinite momentum frame $\gamma \gg 1$, the energy and momentum are the same:

$$P_{(^3\mathrm{He})} = \gamma M_{(^3\mathrm{He})} = 3\gamma M = 3p_i; x = \frac{p_i}{P} = \frac{1}{3}. \tag{5.38}$$

Here, p_i is the momentum carried by the ith constituent, either one of the protons or the neutron, which we have assumed have the same mass. Feynman's x, which plays a crucial role in the parton model, is defined in this way, but in the parton model $0 \leq x \leq 1$. How this comes about was briefly mentioned in Chapter 1.

The parton model is based on the current quarks. They have a small mass, so the mass of the proton is made of kinetic energy, and this is the key difference between the proton and the ^{3}He nucleus. Following the procedure above, we write

$$M = \sum_i E_i, \tag{5.39}$$

where the sum is over all partons in the proton, and E_i is the energy of the ith parton, that is essentially massless. Now, view the proton from the infinite momentum frame as follows:

$$\begin{gathered} P_p = \gamma M = \sum_i \gamma E_i = \sum_i \gamma(E_i + p_{i||}) = \sum_i P_i, \\ x_i = \frac{P_i}{P_p} = \frac{(E_i + p_{i||})}{M}. \end{gathered} \tag{5.40}$$

We can add $\sum_i p_{i||}$ to $\sum_i E_i$ because $\sum_i p_{i||} = 0$, it being a proton rest frame quantity, and the proton is not going anywhere in its rest frame. $p_{i||}$ is a signed quantity, and its maximum value is $p_{i||}^{\max} = M/2$. The maximum value of E_i is the same, so we have

$$0 \leq x_i \leq 1. \tag{5.41}$$

It is perhaps worth noting that no parton is traveling backwards. Partons with x near zero are going backwards in the proton rest frame. In this simple picture, the Feynman x distribution of a quark in the proton is given by the rest frame distributions in energy and parallel momentum. So, if we knew how the quarks are bound, we could calculate the parton distribution function $f(x)$. Perhaps it is worth mentioning that the baryon magnetic moments are not so easy to understand with current quarks. The light quarks would have much larger magnetic moments, several hundred times the proton moment, that have to be canceled out in some way by orbital currents or other effects to add up to μ_p. On the other hand, the constituent quarks would behave like the nucleons in ^{3}He, not like massless partons. So, we have two pictures of the quarks in a baryon, massive and bound for static properties, and light and free for very high energy, that are quite distinct and not easy to reconcile.

Equation (5.22) introduces the two form factors for elastic scattering $F_1(q^2)$ and $F_2(q^2)$ inside a Dirac vector current. For inelastic scattering, although there is a baryon in the final state to conserve baryon number, there are also an accompanying group of other particles, mostly pions, that render the final state spinor $\bar{u}(p')$ unrealistic. We therefore skip the trace stage, and write down the tensor directly in terms of invariants. Call it $W^{\mu\lambda}$, which can depend on $p^\mu, p^\lambda, q^\mu, q^\lambda$, and the metric tensor. As before, we require

$$q_\mu W^{\mu\lambda} = q_\lambda W^{\mu\lambda} = 0. \tag{5.42}$$

Introducing two unknown functions $W_1(\nu, q^2)$ and $W_2(\nu, q^2)$, where for inelastic scattering the two variables $\nu = E - E'$ and q^2 are independent, we write

$$W^{\mu\lambda} = W_1 \left(-g^{\mu\lambda} + \frac{q^\mu q^\lambda}{q^2} \right) + \frac{W_2}{M^2} \left(p^\mu - \frac{p \cdot q}{q^2} q^\mu \right) \left(p^\lambda - \frac{p \cdot q}{q^2} q^\lambda \right). \tag{5.43}$$

The proton mass squared is M^2. It is easy to show that Equation (5.43) satisfies Equation (5.42). This form is Equation (8.27) in [2]. Our treatment of the parton model will follow that of Halzen and Martin, who give a clear presentation of the subject. Equation (5.43) can be contracted with the electron tensor Equation (5.25). The resulting formula is simpler than in the Rosenbluth elastic case because Equation (5.43) is written down after the matrix element has been squared and averaged, and the two structure functions W_1 and W_2 have been judiciously chosen. In order to obtain the form commonly displayed in the literature, it is necessary to multiply by $8\pi M$.

$$\begin{aligned}8\pi M E_{\mu\lambda}W^{\mu\lambda} &= 32\pi M W_1\left(k\cdot k' + \frac{2q\cdot k\; q\cdot k'}{q^2}\right)\\ &\quad + \frac{32\pi M W_2}{M^2}\Bigg\{2k\cdot p\; k'\cdot p - k\cdot k' p^2\\ &\quad - \frac{2p\cdot q}{q^2}(k\cdot q\; k'\cdot p + k\cdot p\; k'\cdot q - k\cdot k'\; q\cdot p)\\ &\quad + \frac{(p\cdot q)^2}{q^4}(2k\cdot q\; k'\cdot q - k\cdot k' q^2)\Bigg\}. \qquad (5.44)\end{aligned}$$

Using the kinematic formulas for the target initially at rest, as in the elastic case, the contraction of the two tensors simplifies. There are several cancellations:

$$8\pi M E_{\mu\lambda}W^{\mu\lambda} = 64\pi M E E'(2W_1\sin^2(\theta/2) + W_2\cos^2(\theta/2)). \qquad (5.45)$$

We seek a differential cross-section for electron scattering in two variables $d\sigma/dE'd\Omega$. The scattering formula that we have been using is specialized to a two-body final state, and that case will not apply to inelastic electron–proton scattering, that has at least a three particle final state $e^- + p \to e^- + n + \pi^+$. At higher $|q^2|$, the number is even greater. A phase space factor for an n particle final state was given in the decay rate formula Equation (4.13) in Chapter 4:

$$(2\pi)^4\delta^4\left(k + p - k' - \sum_{i=1}^{n} p_i'\right)\frac{d^3k'}{(2\pi)^3 2E'}\prod_{i=1}^{n}\frac{d^3p_i'}{(2\pi)^3 2p_i'^0}. \qquad (5.46)$$

For a two-body final state, there are six variables and four delta functions, so two are leftover, that are the two angles θ and ϕ of the scattered beam. For a three-body final state, we learned in Chapter 2 that there are two energies remaining after integration over the delta functions. Actually, there are five remaining variables — two energies plus three Euler angles for the orientation of the decay plane in the center of mass. More final state particles give more independent variables that are required to specify the event. However, in this analysis we do not observe the event multiplicity. We just measure ν and q^2, and integrate over all final states that are consistent with those energy and momentum transfers. All of the phase space and the δ functions have been absorbed in the definition of $W^{\mu\lambda}$. We do however have the flux and the phase space factor for the scattered electron. So we attempt to write down a cross-section as follows:

$$\begin{aligned} d\sigma = {} & \frac{e^4}{4q^4} 64\pi MEE'(2W_1(\nu,q^2)\sin^2(\theta/2) \\ & + W_2(\nu,q^2)\cos^2(\theta/2)) \frac{d^3k'}{(2\pi)^3 2E'} \frac{1}{4ME}. \end{aligned} \tag{5.47}$$

The differential $d^3k' = E'^2 dE' d\Omega$, and $e^4 = 16\pi^2\alpha^2$:

$$\begin{aligned} \frac{d\sigma}{dE'd\Omega} = {} & \frac{\alpha^2}{4E^2\sin^4(\theta/2)} \{2W_1(\nu,q^2)\sin^2(\theta/2) \\ & + W_2(\nu,q^2)\cos^2(\theta/2)\}. \end{aligned} \tag{5.48}$$

In terms of the Mott cross-section, we have

$$\frac{d\sigma}{dE'd\Omega} = \sigma_M \{W_2(\nu,q^2) + 2W_1(\nu,q^2)\tan^2(\theta/2)\}. \tag{5.49}$$

This is the form written down by Bjorken in his classic paper on scaling in electron scattering [5]. The unknown functions W_1 and W_2 are called structure functions, to distinguish from their elastic counterparts, called form factors.

The elastic cross-sections can be written as double differentials with δ functions. For $e^-\mu^- \to e^-\mu^-$,

$$\frac{d\sigma}{d\Omega} = \int dE' \frac{4\alpha^2 E'^2}{q^4} \left\{\cos^2(\theta/2) - \frac{q^2}{2m^2}\sin^2(\theta/2)\right\} \delta(\nu + q^2/2m). \tag{5.50}$$

Here, m is the muon mass, and the delta function is treated in a manner similar to Equations (5.12) and (5.13), as follows:

$$\begin{aligned}
&\int dE'\,\delta(\nu + q^2/2m) \\
&\quad = \int dE'\,\delta\left(E - \frac{2EE'\sin^2(\theta/2)}{m} - E'\right) \\
&\quad = \frac{1}{1 + (2E\sin^2(\theta/2))/m}\int dE'\,\delta\left(E' - \frac{E}{1 + (2E\sin^2(\theta/2))/m}\right) \\
&\quad = \frac{E'}{E}. \qquad (5.51)
\end{aligned}$$

The Rosenbluth formula for elastic electron–proton scattering can be written the same way:

$$\begin{aligned}
\frac{d\sigma}{d\Omega} = \int dE'\,\frac{4\alpha^2 E'^2}{q^4}\left\{\frac{G_E^2 + \tau G_M^2}{1+\tau}\cos^2(\theta/2) + 2\tau G_M^2\sin^2(\theta/2)\right\} \\
\times\,\delta\left(\nu + \frac{q^2}{2M}\right). \qquad (5.52)
\end{aligned}$$

Here, M is the proton mass, and the exercise with the δ function is the same as above.

Richard Taylor from SLAC and Jerome Friedman and Henry Kendall from MIT won the Nobel Prize in physics in 1990 for their experimental work in inelastic electron scattering. Their Nobel addresses give a good accessible overview of that work. The program was a remarkable synergy between theory and experiment, with the important theoretical ideas coming from Richard Feynman and J. D. Bjorken (who is universally known as 'Bj'). Taylor describes the early equipment. First, there was the 20 GeV two-mile long electron linac that supplied the beam of intensity $\sim 100\mu$ amp. Then there was 'End Station A', that housed the 20 and 8 GeV spectrometers used to measure the scattered electron beam. Each spectrometer had several elements — magnetic dipoles, quadrupoles, and correcting sextupoles. To measure a double differential cross-section, the acceptance in momentum and solid angle had to be accurately known. The liquid hydrogen targets were of special design because the high electron beam intensity boiled away the liquid. The electron beam was pulsed by the klystrons that accelerated it, resulting in a 2-μs burst of charge 360 times a second. This poor duty factor made it impossible at that time to measure the recoils in coincidence with the scattered electrons, so that the experiments only measured the final state electron, and integrated over all hadronic recoils that corresponded to a given ν and q^2.

In his paper in the *Physical Review* noted above, Bjorken argued that as the variables ν and $|q^2|$ went to infinity, keeping the ratio $x = |q^2|/2M\nu$ fixed, the structure functions W_1 and W_2 should become functions of x alone, and not ν and q^2 separately. Two new structure functions were introduced (not to be confused with the proton form factors in elastic scattering) defined as

$$MW_1(\nu, q^2) \to F_1(x),\ \nu W_2(\nu, q^2) \to F_2(x),\ x = -\frac{q^2}{2M\nu}, \qquad (5.53)$$

where M is the proton mass, $\nu = E - E' = p \cdot q/M$, and the four vector $q = k - k' = p' - p$. This is called 'scaling'. I believe that it is safe to say that very few people understood Bjorken's reasoning, but an experimental test of the idea was relatively straightforward. Plot a measured proton structure function, say $W_1(\nu, q^2)$, as a function of $|q^2|$ for fixed x. If scaling holds, the plot would be flat, indicating that $MW_1(\nu, q^2) \to F_1(x)$. This was done with early inelastic scattering data, and the curve was flat! It was not necessary for $-q^2 \to \infty$; in fact, scaling seemed to work for $-q^2 > 1\,\mathrm{GeV}^2$. Richard Feynman visited SLAC about this time (1969). He had been thinking about hadron production, and was developing what he called the parton model (Feynman coined the term parton) of pointlike constituents of protons and pions that carried a fraction $x_i = p_i/P$ of the hadron momentum as viewed from an infinite momentum frame. It would take a while for people to realize that Feynman's x and Bjorken's x were the same, and that Feynman's partons were fractionally charged quarks and uncharged gluons.

The 1990 Nobel Laureates state that when they began their experimental program to measure what is now called 'deep inelastic scattering' of very high energy electrons off of hydrogen and deuterium targets, they had no idea that they were about to demonstrate that the proton is composed of massless spin 1/2, one-third integer charge quarks, and massless uncharged spin 1 gluons. It took more than ten years, but that is exactly what they did, with Feynman, Bjorken, and other theorists looking over their shoulder. The Nobel Prize is fickle. Gell-Mann won it for quarks, but not George Zweig, who independently proposed the same thing. Cabibbo was omitted from the award to Kobayashi and Maskawa, for what everyone calls the CKM matrix. Feynman won it for quantum electrodynamics, sharing it with Schwinger and Tomonaga. This was certainly well deserved, but Feynman made many contributions to physics, and merited the prize all by himself. Friedman, Kendall, and Taylor were the leaders of a group that did the very accurate measurements necessary to convince doubters of the quark–parton model. Awards to leaders of large groups have been more frequent in recent years.

More accurate data have shown that scaling does not precisely hold. Some very pretty data from the CDHS experiment at CERN [6], which studied neutrino scattering in iron, are shown in Figure 5.4. We shall soon discuss neutrino scattering in the parton model. Neutrinos have some advantages over electrons. The interaction does not depend on the quark charge, and $\nu_\mu + d \to \mu^- + u$, $\bar{\nu}_\mu + u \to \mu^+ + d$, but not with the quarks interchanged, so neutrinos sample the down quark distribution, and anti-neutrinos sample the up quark. There are also interesting spin effects with the weak interaction. Quarks, negative muons, and muon neutrinos are all left-handed, but the muon anti-neutrino and the positive muon are right-handed, so the angular distributions of the final state muons would be different for the two reactions. The figure clearly shows that $F_2(x)$ is also a function of $Q^2 = -q^2$. At small x, F_2 increases with increasing Q^2, and at large x, it decreases with increasing Q^2. The structure function $F_2(x)$ shifts to lower x at high Q^2. This effect can be explained by the Altarelli–Parisi equations, displayed but not derived later in this chapter.

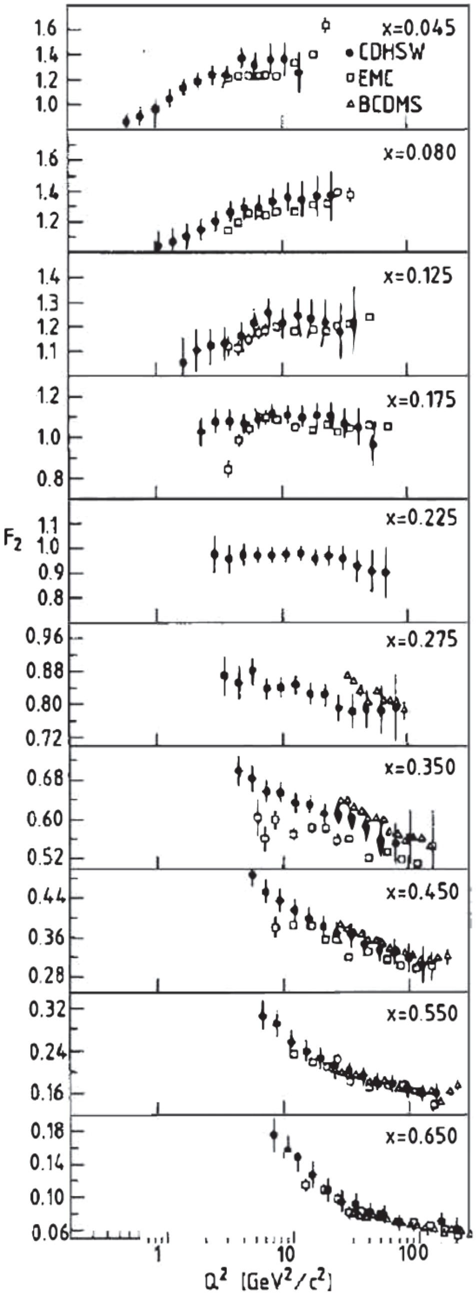


Figure 5.4: (*Continued*)

Figure 5.4: The proton structure function $F_2(x, Q^2)$ plotted as a function of Q^2 in bins of increasing x from top to bottom. The x-axis is logarithmic, but the y-axis is linear. The range changes to reflect the x dependence of F_2. The solid dots are neutrino data. The open points are muon data, multiplied by 18/5, a quark model charge factor for a target with equal numbers of protons and neutrons $\langle q^2\rangle = (q_u^2 + q_d^2)/2 = 5/18$. The neutrinos do not couple to the quark charges. The scaling violation is clearly seen. At small x, F_2 increases with Q^2, and at large x, F_2 decreases. Credit Ref. [6].

Feynman was using some empirical features of hadron production in formulating his parton model. Multiple pion production had been studied at high energy proton machines, with the following approximate rules:

1. The pion transverse momentum distribution falls off exponentially with increasing $p_T = p\sin\theta$.
$$\frac{d\sigma}{dp_T} \sim \sigma_0 e^{-p_T/p_{T0}}, \tag{5.54}$$
where p_{T0} increases slowly with energy, and is typically about 0.4 GeV/c at $\sqrt{s} = 1000\,\mathrm{GeV}$. Hence, even at this rather large value of $\sqrt{s}$ (the Fermilab Tevatron collider) only about 1% of the pions have $p_T > 2\,\mathrm{GeV/c}$.
2. The average pion multiplicity increases only logarithmically with increasing energy.
3. Given items 1 and 2, as the energy available increases, the only place it can go is into the parallel component of the pion momentum.

There is a nice argument for the multiplicity. The pion rapidity is defined as

$$y = \frac{1}{2}\ln\frac{E + p_{||}}{E - p_{||}};\ E_{\max} = p_{||\max} = \sqrt{s}/2;\ y_{\max} = \ln\big(\sqrt{s}/m_\pi\big). \tag{5.55}$$

The factor m_π in the denominator comes from expanding $E_{\max}$. So, think of a box centered at $y = 0$ and extending $\pm y_{max}$ in the center of mass. This length grows with $\ln(\sqrt{s})$. For the Tevatron collider at $\sqrt{s} = 2000\,\mathrm{GeV}$, $\pm y_{\max} = 19$ units of rapidity. If the number of pions per unit rapidity dn/dy is constant, then the multiplicity grows with $\ln(\sqrt{s})$. It is like teeth in a comb — for constant spacing, the longer the comb, the more teeth. This idea only approximately works. At 7 GeV, $dn/dy \sim 2$, and at 7000 GeV (the LHC), $dn/dy \sim 10$. So, it is not really constant, but we have increased the energy by a factor of 1000, and the multiplicity per unit rapidity has only gone up a factor of five. This can be explained by adding a term $\sim (\ln\sqrt{s})^2$ to dn/dy.

Another rule of thumb well above energy thresholds is that strange mesons, K′s and $\bar{\mathrm{K}}'s$, occur at about the 10% level, independent of the energy. Baryons are at the few percent level. This does not mean that the recoil hadrons in deep electron scattering do not always include a nucleon. It means rather that nucleon–anti-nucleon pairs occur among the mesons a few percent of the time. Mesons that contain the heavier quarks, D's and B's, are below 1%. The rules are satisfied by any hadronic system, including the unobserved recoil of mass $W = \sqrt{(p+q)^2}$. The limitation in transverse momentum, item #1 above, appears in what is called the fragmentation of the leading quark into hadrons after the hard scattering. The outgoing quark forms a 'jet' of particles in a cone, the axis of which is approximately the original quark direction. This process is verified in the annihilation $e^-e^+ \to q\bar{q}$, where the axes of the opposite jets balance, and the angular distribution is the same as $e^-e^+ \to \mu^-\mu^+$. Small transverse momentum and large longitudinal momentum were components of the parton model of the proton discussed above.

In our first-order theory of deep inelastic scattering, the virtual photon interacts with the proton. The virtual photon is a bit peculiar, in that it has negative invariant mass squared — it is a space like four vector $q^2 = (k-k')^2 = (E-E')^2 - (\vec{k}-\vec{k}')^2 = -2EE'(1-\cos\theta)$. The z axis is normally chosen along the incident electron direction, but for consideration of the virtual photon and its polarization, choose the z axis along $\vec{q} = \vec{k} - \vec{k}'$. Then

$$q^\mu = (\nu, 0, 0, |\vec{k} - \vec{k}'|),$$

and the polarization vectors are

$$\epsilon^{\mu}_{\lambda=\pm 1} = \mp\frac{1}{\sqrt{2}}(0,1,\pm i,0); \quad \text{and} \quad \epsilon^{\mu}_{\lambda=0} = \frac{1}{\sqrt{-q^2}}(|\vec{k}-\vec{k}'|,0,0,\nu). \tag{5.56}$$

Note that

$$\epsilon^{\mu}_{\lambda} q_{\mu} = 0; \quad \epsilon^{\mu *}_{\lambda}\epsilon_{\mu,\lambda'} = 0, \lambda \neq \lambda'; \quad \epsilon^{\mu *}_{\lambda}\epsilon_{\mu\lambda} = -1; \quad \lambda = \pm 1; \quad \text{and}$$
$$\epsilon^{\mu}_{\lambda}\epsilon_{\mu\lambda} = +1; \quad \lambda = 0. \tag{5.57}$$

So, the longitudinal component, that is absent for a real photon, is a time-like vector, while the transverse components are space-like. All components are orthogonal to q^{μ}, but unlike a real photon, $q^{\mu}q_{\mu} < 0$. Recall that for a massive particle the spin direction is defined in the rest frame, where it has only space components, so it is space-like, and orthogonal to the four momentum, which has only a time component in the rest frame. Although $q^2 \neq 0$, the virtual photon has no rest frame, because its 'velocity' exceeds unity. Time-like virtual photons also exist, like the intermediate state in $e^-e^+ \to \mu^-\mu^+$. In the laboratory frame of an e^-e^+ storage ring collider, the intermediate photon has only energy, and no momentum, so it is at rest!

The completeness relation for the virtual photon polarization is

$$\sum_{\lambda}(-1)^{\lambda+1}\epsilon^{\mu *}_{\lambda}\epsilon^{\nu}_{\lambda} = -g^{\mu\nu} + \frac{q^{\mu}q^{\nu}}{q^2}, \tag{5.58}$$

where the sign changes for the time-like component $\lambda = 0$. We argued in Chapter 4 that the completeness relation for intermediate photons was $\sum_{\lambda}\epsilon^{\mu *}_{\lambda}\epsilon^{\nu}_{\lambda} \to -g^{\mu\nu}$. It is easy to show with the spinors and the Dirac equation that the extra term in Equation (5.58) does not couple to a vector current $\bar{u}(k')\gamma^{\mu}u(k)$.

The cross-section for a virtual photon with polarization λ interacting with the target proton to create the multiparticle final state is linear in the fine structure constant $\alpha = e^2/(4\pi)$, and is usually written as

$$\sigma_\lambda = \frac{4\pi^2\alpha}{K}\epsilon_\lambda^{\mu *}\epsilon_\lambda^{\nu}W_{\mu\nu} = \frac{4\pi^2\alpha}{K}\epsilon_\lambda^{\mu *}\epsilon_\lambda^{\nu}\left\{W_1\left(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2}\right)\right.$$
$$\left. + \frac{W_2}{M^2}\left(p_\mu - \frac{p\cdot q}{q^2}q_\mu\right)\left(p_\nu - \frac{p\cdot q}{q^2}q_\nu\right)\right\}, \tag{5.59}$$

where we have used Equation (5.43), and introduced the new variable $K = (W^2 - M^2)/(2M)$. If the photon were real, the four vector $q^2 = 0$, and K is the photon energy that enters into the flux factor. The flux for a virtual photon is taken as the same expression, so that for the virtual case

$$K = \frac{W^2 - M^2}{2M} = E - E' + \frac{q^2}{2M}. \tag{5.60}$$

The z axis is along $\vec{k} - \vec{k}'$, so the momentum four vectors are

$$p = (M, 0, 0, 0);\ q = (E - E', 0, 0, |\vec{k} - \vec{k}'|). \tag{5.61}$$

The angle θ is between $\vec{k}$ and $\vec{k}'$, and is independent of the choice of the z axis. For $\lambda = \pm 1$, the four vectors are orthogonal: $\epsilon\cdot p = \epsilon\cdot q = 0$. So,

$$\sigma_{\lambda=\pm 1} = \frac{4\pi^2\alpha}{K}W_1. \tag{5.62}$$

For $\lambda = 0$, $\epsilon\cdot q = 0$, and $\epsilon\cdot p = M|\vec{k} - \vec{k}'|/\sqrt{-q^2}$, giving

$$\sigma_{\lambda=0} = \frac{4\pi^2\alpha}{K}\left(-W_1 + W_2\frac{(\vec{k} - \vec{k}')^2}{q^2}\right)$$
$$= \frac{4\pi^2\alpha}{K}\left(-W_1 + W_2\left(1 - \frac{(E - E')^2}{q^2}\right)\right). \tag{5.63}$$

Equation (5.63) must vanish as $q^2 \to 0$. This identification shows that two independent structure functions are needed for the relative contributions of the transverse and longitudinal polarizations of the virtual photon.

Equation (5.48) can be used to compare the two ways of writing the deep inelastic cross-section:

$$\frac{d\sigma}{dE' d\Omega} = \frac{4\alpha^2 E'^2}{q^4}\left\{2W_1 \sin^2(\theta/2) + W_2 \cos^2(\theta/2)\right\}$$
$$= \frac{4\pi^2\alpha}{K}\left\{(A-B)W_1 + BW_2\left(1 - \frac{\nu^2}{q^2}\right)\right\}, \tag{5.64}$$

where $\sigma_T = (4\pi^2\alpha)/K \times A$, and $\sigma_L = (4\pi^2\alpha)/K \times B$. A little algebra gives

$$A = \frac{K\alpha E'}{2\pi^2 E|q^2|}\frac{|q^2|\cos^2(\theta/2) + 2\sin^2(\theta/2)(\nu^2 + |q^2|)}{2\sin^2(\theta/2)(\nu^2 + |q^2|)}, \tag{5.65}$$

and

$$B = \frac{K\alpha E'}{2\pi^2 E|q^2|}\frac{|q^2|\cos^2(\theta/2)}{2\sin^2(\theta/2)(\nu^2 + |q^2|)}. \tag{5.66}$$

The ratio $B/(A + B)$ is the fraction of longitudinal polarization of the virtual photon.

5.5. The Parton Model

We now have the materials to derive an expression for deep inelastic scattering $e^- + p \to e^- + X$ in terms of elastic scattering of the electron off a Dirac point-like quark inside the proton. After elastic scattering, the recoil quark leaves the spectator partons behind. Then on a longer timescale than the hard elastic collision, all of the leftover debris plus the struck quark assemble themselves into the hadronic final state that has no color. Just how this takes place does not concern us at the present time. We are interested in formulating the elastic collision. To get started, we rewrite Equation (5.50) for e^-q elastic scattering without the integral over dE', as follows:

$$\frac{d\sigma}{d\Omega dE'} = \frac{4e_i^2\alpha^2 E'^2}{q^4}\left\{\cos^2(\theta/2) - \frac{q^2}{2m^2}\sin^2(\theta/2)\right\}\delta\left(\nu + \frac{q^2}{2m}\right), \tag{5.67}$$

where m is the quark mass, and e_i is the quark charge in units of the elementary charge. Since the parton model involves the ‘infinite momentum' frame, it is useful to write the cross-section in terms of only the invariant variables s, t, and u. We have the following relations:

$$s = (p+k)^2 = m^2 + 2p\cdot k = m^2 + 2mE;\; t = (k-k')^2 = q^2;$$
$$u = (p-k')^2 = m^2 - 2p\cdot k' = m^2 - 2mE';$$
$$2m\nu = 2m(E-E') = 2p\cdot(k-k') = s+u-2m^2. \tag{5.68}$$

Now, for some manipulations with the δ function

$$\delta(\nu + t/2m) = 2m\delta(2m\nu + t) = 2m\delta(s+t+u-2m^2). \tag{5.69}$$

The argument of this δ function should look familiar. To change the variables of integration, we need the Jacobian. The argument proceeds as follows:

$$t = -2EE'(1-\cos\theta),\; u = m^2 - 2mE',$$
$$\partial t/\partial E' = -2E(1-\cos\theta),\; \partial t/\partial\cos\theta = 2EE',$$
$$\partial u/\partial E' = -2m,\; \partial u/\partial\cos\theta = 0,\; J = 4mEE',$$
$$2\pi dt du = 4mEE' dE' d\Omega. \tag{5.70}$$

Converting the angle functions into the invariant variables gives

$$\frac{d\sigma}{dtdu} = \frac{4\pi e_i^2\alpha^2}{t^2(s-m^2)^2}\left\{(s-m^2)(m^2-u) + m^2 t + t^2/2\right\}$$
$$\times\,\delta(s+t+u-2m^2). \tag{5.71}$$

In the limit that we can ignore the quark mass, this formula becomes

$$\frac{d\sigma}{dtdu} = \frac{2\pi e_i^2\alpha^2}{t^2}\left\{\frac{s^2+u^2}{s^2}\right\}\delta(s+t+u). \tag{5.72}$$

Now, introduce the 'hat' variables that apply to the parton, using Feynman's $x = p_i/p$, and ignoring all masses, including that of the proton:

$$s = (p+k)^2 = 2p\cdot k; \quad \hat{s} = xs; \quad t = (k-k')^2 = \hat{t};$$
$$u = (p-k')^2 = -2p\cdot k'; \quad \hat{u} = xu. \tag{5.73}$$

For the scattered electron, $\hat{t} = t$, while s and u are decreased by the momentum fraction x. The cross-section in terms of the hat variables then is

$$\frac{1}{x}\frac{d\sigma}{dtdu} = \frac{d\sigma}{d\hat{t}d\hat{u}} = \frac{2\pi e_i^2\alpha^2}{\hat{t}^2}\left\{\frac{\hat{s}^2+\hat{u}^2}{\hat{s}^2}\right\}\delta(t+x(s+u)). \tag{5.74}$$

Note that the expression in the curly brackets is invariant under the substitution $s \to xs$ and $u \to xu$, so that Equation (5.74) can be written as a function of x in terms of laboratory variables

$$\frac{d\sigma}{dtdu} = \frac{2\pi x e_i^2\alpha^2}{t^2}\left\{\frac{s^2+u^2}{s^2}\right\}\delta(t+x(s+u)). \tag{5.75}$$

The argument of the δ function is

$$\text{Feynman } x = -\frac{t}{s+u} = -\frac{q^2}{2M(E-E')} = -\frac{q^2}{2M\nu} = x \text{ Bjorken,} \tag{5.76}$$

where we have re-introduced the proton mass to recover Bjorken's form. This important relation shows that the two x's are the same. Bjorken's x is the fraction of the proton momentum carried by the struck quark. The intermediate virtual photon picks out just this quark to hit. The parton model then asserts that the inelastic cross-section can be written as a sum over scattering by individual quarks carrying momentum fraction x inside the proton:

$$\frac{d\sigma(e^-p\to e^-X)}{dtdu} = \sum_i \int_0^1 dx f_i(x)\frac{d\sigma(e^-q_i\to e^-q_i)}{dtdu}. \tag{5.77}$$

Here, $f_i(x)$ is the quark distribution function, and the sum is over all charged partons in the proton. The gluons carry no charge, and do not couple to the virtual photon, but do carry momentum, so the closure relation pertains to all partons, not just the quarks:

$$\sum_{\text{all partons}} \int_0^1 dx\, x f_i(x) = 1. \tag{5.78}$$

Now, returning to Equations (5.48) and (5.53), and substituting (5.53) into (5.48), we get:

$$\frac{d\sigma}{dtdu} = \frac{2\pi\alpha^2}{4E^2\sin^4(\theta/2)}\frac{1}{4MEE'}\left\{\frac{2}{M}F_1(x)\sin^2(\theta/2)\right.$$
$$\left. + \frac{1}{\nu}F_2(x)\cos^2(\theta/2)\right\}. \tag{5.79}$$

The parton model requires Equation (5.79), when written in terms of the Mandelstam variables, and Equation (5.77) to be the same. Using the formulas in Equation (5.68), with the mass equal to the proton mass, and in the end dropping terms proportional to M, we obtain the following form for Equation (5.79):

$$\frac{d\sigma}{dtdu} = \frac{4\pi\alpha^2}{s^2t^2}\frac{1}{s+u}\left\{(s+u)^2xF_1(x) - usF_2(x)\right\}. \tag{5.80}$$

Now, inserting Equation (5.75) into Equation (5.77), we get:

$$\frac{d\sigma}{dtdu} = \sum_i \int_0^1 dx f_i(x)\left\{\frac{2\pi x e_i^2\alpha^2}{s^2t^2}(s^2+u^2)\delta(t+x(s+u))\right\}. \tag{5.81}$$

Evaluating the δ function integral — that gives the factor $1/(s+u)$ — and equating the two forms for the cross-section gives

$$\frac{2\pi\alpha^2(s^2+u^2)}{s^2t^2}\frac{1}{s+u}\sum_i e_i^2 x f_i(x)$$
$$= \frac{4\pi\alpha^2}{s^2t^2}\frac{1}{s+u}\left\{(s+u)^2xF_1(x) - usF_2(x)\right\}. \tag{5.82}$$

This is what Halzen and Martin call the master formula for the parton model, namely

$$2xF_1(x) = F_2(x) = \sum_i e_i^2 x f_i(x). \tag{5.83}$$

Another dimensionless invariant variable called y, not to be confused with the rapidity, is often introduced:

$$y = \frac{2p \cdot q}{2p \cdot k} = \frac{2p \cdot k - 2p \cdot k'}{s} = \frac{s+u}{s} = \frac{\hat{s}+\hat{u}}{\hat{s}};$$

$$-\frac{\hat{u}}{\hat{s}} = 1 - y; \quad 0 \leq y \leq 1; \;\; |q^2| = xys. \tag{5.84}$$

To change the variables, $-t = xys$; $u = s(y-1)$, so $dtdu = ys^2dxdy$, and

$$\frac{d\sigma}{dxdy} = \frac{2\pi\alpha^2}{t^2} s(1 + (1-y)^2)F_2(x), \tag{5.85}$$

where again

$$F_2(x) = \sum_i e_i^2 x f_i(x).$$

Since s/t^2 has dimensions $1/E^2$ that can be converted to a cross-section via the relation $1\text{GeV}^{-2} = 0.389$ mb, the rest of the formula, including the parton distribution functions, is dimensionless. The factor $(1 + (1-y)^2)$ in Equation (5.85) comes from the formula for $e^-\mu^-$ scattering, and confirms that the partons have spin 1/2.

5.6. Parton Distribution Functions for Protons and Neutrons

The parton distribution functions, $f_i(x)$ in Equation (5.83), are determined by fits to experimental data. There are several groups that work on these fits. The CTEQ collaboration[a] has done the fits shown in Figure 5.5. A large amount of data goes in to obtaining the individual quark and gluon structure functions. A part of these data, deep inelastic e^-p collisions from the HERA electron/positron proton colliding rings mentioned earlier, is shown in Figure 5.6. The data are shown as points with error flags, and the CTEQ fits are the curves. The so-called scaling violation, the $|q^2|$ dependence at fixed

---

[a] See the CTEQ web site at physics.smu.edu/scalise/cteq. Codes for calculating structure functions are available there.

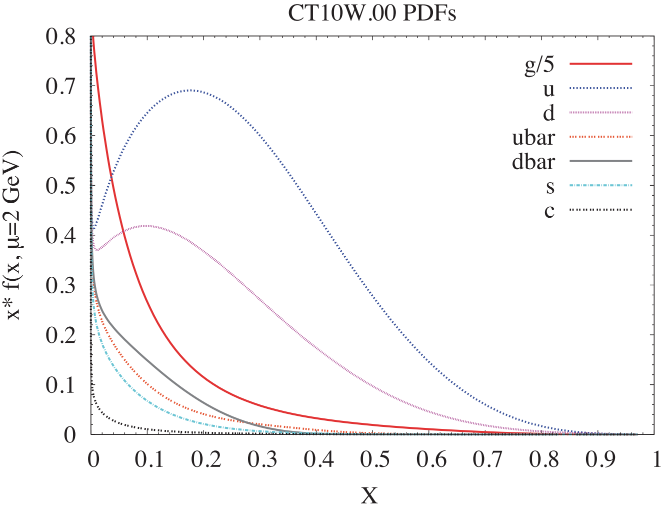


Figure 5.5: Parton distribution functions for the proton, calculated by the CTEQ collaboration. The gluon distribution is peaked a low x. The u and d distributions are valence plus sea. The $\bar{u}, \bar{d}, s$, and c are all sea distributions. The valence quark distributions are $u_v(x) = u(x) - \bar{u}(x)$, and $d_v(x) = d(x) - \bar{d}(x)$. The subtraction removes some of the low x distribution. The parameter μ is an energy in QCD that appears in the calculations.

x, is rather modest in this high $|q^2|$ dataset. These HERA data do not extend past $x = 0.65$. The remainder of this section is devoted to a description of how the data are untangled to obtain the parton distributions shown in Figure 5.5.

Equation (5.83) is a sum over all quark distributions in the proton, weighted by the square of their fractional charges, yet Figure 5.5 shows individual quarks. How did they sort it out? Also, where did the gluon distribution come from, since gluons do not couple to the virtual photon? Although deep inelastic scattering of electrons or neutrinos by nucleons has been historically the principal source of data on the parton distributions, they also play a role in purely hadronic collisions, and data for many hadronic processes are a main source for the gluon distribution. Modern fits use all of the relevant experimental data that exists.

Hadronic hard collisions, as occur for example in a $p\bar{p}$ storage ring with a large exchange of energy, take place in the parton model as an

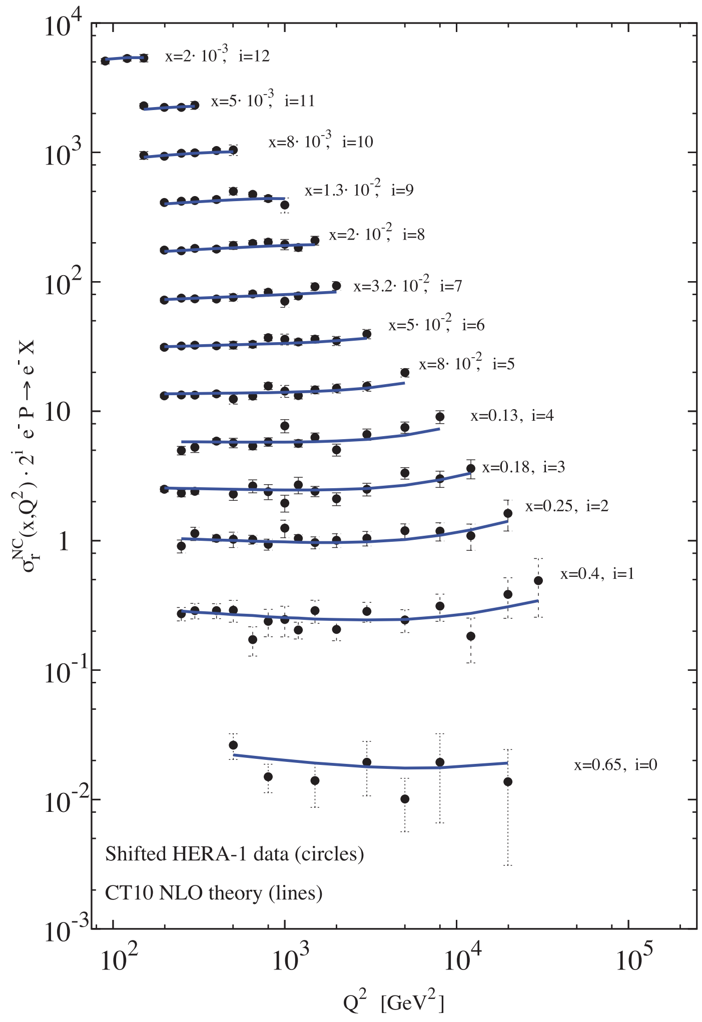


Figure 5.6: Experimental data from the HERA electron–proton colliding rings for $F_2(x, |q^2|)$ showing the CTEQ fits.

interaction between individual partons in the two hadrons, with the rest of the hadronic material being spectators. A relatively simple reaction that can be described by the parton model is the production of high mass lepton pairs in hadronic collisions, for example, $p + \bar{p} \to \mu^- \mu^+ + X$. The parton model describes this process as

$q\bar{q}$ annihilation to produce a muon pair via an intermediate photon, in analogy with $e^-e^+ \to \mu^-\mu^+$. This is called the Drell–Yan mechanism, after the theorists who originally proposed it [7]. It accounts for the production of high mass pairs, in contrast with the Bethe–Heitler mechanism, in which an incident charged particle radiates a virtual photon that creates a lepton pair (e^-e^+ or $\mu^-\mu^+$), and one of the members of the final state pair scatters off the nucleus to conserve four momentum. This process creates low mass pairs, as the cross-section $\sigma \sim 1/\omega^3$, where ω is the energy of the intermediate photon. The Drell–Yan process, on the other hand, creates high mass lepton pairs, and also the massive intermediate vector bosons $W^\pm$ and Z. For the reaction $p+\bar{p} \to \mu^-\mu^+ + X$, u quarks in the proton annihilate with $\bar{u}$ quarks in the antiproton. The distributions of quarks in the proton and antiquarks in the antiproton are assumed identical, so if the proton quark distributions are $u(x)$ and $d(x)$, then muon pair production is proportional to $(4u(x_1)u(x_2) + d(x_1)d(x_2))/9$; where x_1 is in the proton, and x_2 is in the antiproton, and the numerical factors are the squares of the charges. The di-muon invariant mass squared is $m^2_{\mu\mu} = \hat{s} = x_1x_2s$, and the momentum fraction of the muon pair is $(x_1 - x_2)$ in the proton direction. In a $\bar{p}p$ single ring collider, the total momentum is zero, and the muon pair can move in either direction. The annihilations $q\bar{q} \to \gamma$ or $q\bar{q} \to W^+$ are electroweak processes, and hence have smaller cross-sections than purely hadronic reactions like $q\bar{q} \to g$, where g is an intermediate gluon. The purely hadronic reaction $q\bar{q} \to q'\bar{q}'$ produces two back-to-back jets from the hadronization of the outgoing quarks. About 2/3 of all W's decay into quarks as well: $W^+ \to u\bar{d}$, and would show a mass peak amid the hadronic background if the mass resolution from jet energy measurements were good enough. Unfortunately, the jet mass resolution is only marginally adequate, but there is a source of W's for which the hadronic background is small, and that is daughter W's from top quark decay: $t \to W^+ + b$. Top quarks are produced in pairs, and if leptonic decay is used to tag one of the W's, the other one can be reconstructed from di-jet decay. The top production cross-section at the LHC is large enough to furnish a statistically significant sample of $W \to q\bar{q}$, and hence give a direct measurement of the mass resolution around 100 GeV from di-jets. The answer is a full width of about 20 GeV.

Pions also contain quarks, and although the pion structure functions are less well known, high mass di-muon production by pions may be described by Drell–Yan mechanism. Pions, like anti-protons, have valence antiquarks: $\pi^+ = (u\bar{d})$, and $\pi^- = (d\bar{u})$. On a carbon target, with equal numbers of protons and neutrons, because the charge ratio $(u\bar{u}/d\bar{d}) = 4$, we would expect more muon pairs from π^- than π^+. Experimental data [8] shows the ratio of muon pairs from 225 GeV/c $\pi^\pm$ on a carbon target in the di-muon mass range $0.7 \leq m_{\mu\mu} \leq 3.7$ GeV to slowly drop from unity to 1/2. The parton model would be confirmed by a ratio reaching 0.25 at high di-muon mass. A Bethe–Heitler type process would have no preference for π^- compared to π^+.

Much information has been obtained from a comparison of proton and neutron deep inelastic scattering data. The early neutron data came from deuterium targets, from which one obtains approximately the incoherent sum of neutron and proton scattering. Since proton scattering can be independently determined from hydrogen, the neutron cross-section, apart from corrections due to binding, is the difference. The removal of all of the binding effects is complicated, but we will assume that it was successful. The valence quarks are $p = (uud)$, and $n = (ddu)$. The u and d quarks are the only ones that have non-zero I-spin. I-spin invariance requires the proton distribution of u quarks to be equal to the neutron distribution of d quarks. So, in terms of valence and sea quarks, we may write:

$$F_2^{ep}(x) = \frac{x(4u_v(x) + d_v(x))}{9} + xS(x),$$

$$F_2^{en}(x) = \frac{x(4d_v(x) + u_v(x))}{9} + xS(x),$$

$$F_2^{ep} - F_2^{en} = \frac{x(u_v(x) - d_v(x))}{3},$$

$$S(x) = \frac{4}{9}(u_s(x) + \bar{u}(x)) + \frac{1}{9}(d_s(x) + \bar{d}(x) + s(x) + \bar{s}(x)). \quad (5.86)$$

The sea quarks, given by $S(x)$, are virtual $q\bar{q}$ pairs, created by the gluons, and clump at low x. Quark–antiquark pairs carry no proton quantum numbers, and one assumes that $\bar{u}(x) = u_s(x)$ and $\bar{d}(x) = d_s(x)$, so one may write

$$S(x) = \frac{8}{9}\bar{u}(x) + \frac{2}{9}(\bar{d}(x) + \bar{s}(x)). \tag{5.87}$$

The gluon distributions, which are the sources for the sea quarks, are assumed to be the same for the proton and the neutron. $s(x)$ is the strange quark distribution, which is all sea, since the proton has no valence strange quarks. The sea quarks cancel in the difference between the two F_2's. The u quark distribution shown in the figure is $u_v + u_s$. Earlier fits assumed that the strange quark distribution would be the same as $\bar{u}$, but a close inspection of the figure shows that this is not the case. The heavier strange quark is suppressed. The charmed quark, with a mass greater than the mass of the proton, is almost invisible. We left it out of the sum over all sea quarks. There is no sign of b quarks at all. Past $x \geq 0.2$, $4u(x)/9$ is much larger than $S(x)$, so one can solve for $u(x)$ and $d(x)$ from the the two F_2's:

$$x(u(x) + \bar{u}(x)) = \frac{9}{15}(4F_2^{ep}(x) - F_2^{en}(x)) - \frac{2}{5}x\bar{s}(x), \tag{5.88}$$

and

$$x(d(x) + \bar{d}(x)) = \frac{9}{15}(4F_2^{en}(x) - F_2^{ep}(x)) - \frac{2}{5}x\bar{s}(x). \tag{5.89}$$

Here, $u(x) = u_v(x) + u_s(x)$, and $u_s(x) = \bar{u}(x) = d_s(x) = \bar{d}(x)$. To the extent that one can ignore the strange quarks, these equations gives the proton u and d quark–parton distributions in terms of directly measured structure functions.

The valence quarks are normalized as follows:

$$\int_0^1 dx(u(x) - \bar{u}(x)) = 2; \quad \int_0^1 dx(d(x) - \bar{d}(x)) = 1;$$
$$\int_0^1 dx(s(x) - \bar{s}(x)) = 0. \tag{5.90}$$

It is $xu(x)$, the average momentum fraction, that is measured, rather than $u(x)$ alone. The average momentum fraction for all the partons must be unity:

$$\int_0^1 dx\; x(u(x) + d(x) + \bar{u}(x) + \bar{d}(x) + s(x) + \bar{s}(x) + g(x)) = 1. \tag{5.91}$$

Adding Equations (5.88) and (5.89) and integrating gives

$$\begin{aligned}&\int_0^1 dx\; x(u(x) + \bar{u}(x) + d(x) + \bar{d}(x))\\ &\quad = \frac{9}{5}\int_0^1 dx(F_2^{ep}(x) + F_2^{en}(x)) - \frac{4}{5}\int_0^1 dx x\bar{s}(x).\end{aligned} \tag{5.92}$$

The proton and neutron structure functions have been measured, and can be numerically integrated as follows:

$$\int_0^1 dx F_2^{ep}(x) = 0.18; \quad \int_0^1 dx F_2^{en}(x) = 0.12. \tag{5.93}$$

So, dropping the strange quark term gives the following integral:

$$\int_0^1 dx\; x(u(x) + \bar{u}(x) + d(x) + \bar{d}(x)) = 0.54. \tag{5.94}$$

The conclusion is that about half of the total momentum of the proton is carried by the gluons. This argument gives, at least approximately, $\int_0^1 dx x g(x)$, but does not tell us what the shape is, and Figure 5.5 shows $g(x)$. The shape comes from at least two sources. One source is the study of high energy hadronic reactions. Jet production in $p\bar{p}$ collisions at 2 TeV (Fermilab Tevatron) is dominated by gluon–gluon scattering. Another source is theoretical, called the evolution equations, which connect the quark pair production, and hence the sea quarks, to the gluon distribution using perturbative QCD. The derivation of these equations is beyond the scope of this text, but writing them down is perhaps worthwhile, to show how the theory is connected. These equations are given in [9, p. 590]. They are written in terms of $Q^2 = -q^2$, a popular definition in deep inelastic scattering. Q^2 is positive, and Q is real. The derivatives are with respect to $d \ln Q = dQ/Q$. They are named the Altarelli–Parisi equations, after the theorists who first derived them. There are three of them, but one is sufficient to show what is involved. A quark can radiate a gluon, forming a $q+g$ final state; gluons can create $q\bar{q}$ pairs, and $q\bar{q}$ pairs annihilate to form gluons. Additionally, because of the nonlinearity of QCD — the gluons carry color — a gluon can split into two gluons. The various parton distributions must be in equilibrium with all of this going on. Perturbative QCD gives relations among the parton distributions that are proportional to the strong coupling constant $\alpha_s(Q^2)$, defined by the following equation:

$$\alpha_s(Q^2) = \frac{2\pi}{(11 - 2n_f/3)\ln(Q/\Lambda)}. \tag{5.95}$$

Here, n_f is the number of quark flavors above threshold Q. For the light quarks $n_f = 3$, and at the Z mass $Q = 91\,\text{GeV}$, $n_f = 5$, (u, d, s, c, b). The cut-off Λ is the energy at which $\alpha_s \to \infty$, normally chosen around $\Lambda = 200\,\text{MeV}$. Perturbation expansions in powers of α_s are feasible only if $Q \gg \Lambda$. For $\Lambda = 200\,\text{MeV}$ and $n_f = 5$, $Q = M_Z = 91\,\text{GeV}$, a little arithmetic gives $\alpha_s(M_Z) = 0.13$. This is more than a factor of ten larger than the fine structure constant α.

The Altarelli–Parisi equation for the gluon distribution function is

$$\frac{df_g(x,Q)}{d\ln Q} = \frac{\alpha_s(Q^2)}{\pi}\int_x^1 \frac{dz}{z}\Bigg\{P_{g\leftarrow q}(z)\sum_f [f_f(x/z,Q) + f_{\bar{f}}(x/z,Q)] + P_{g\leftarrow g}(z) f_g(x/z,Q)\Bigg\}. \tag{5.96}$$

That the equation is proportional to $\alpha_s(Q^2)$ means that it is a radiative correction, and since at reasonable Q^2 $\alpha_s < 1$, the sensitivity to the gluon structure function is attenuated. The functions $P_{g\leftarrow q}$ and $P_{g\leftarrow g}$ are called splitting functions. $P_{g\leftarrow q}(z)$ gives the probability that a q or a $\bar{q}$ radiates a collinear gluon with momentum fraction

z, leaving the quark/antiquark with $(1-z)$. This splitting function multiplies each of the quark and antiquark parton distribution functions, and adds low x gluons to $f_g(x,Q)$. Similarly, $P_{g\leftarrow g}(z)$ gives the probability that a gluon will radiate another collinear gluon, leaving a pair of gluons with momentum fractions z and $(1-z)$. The three gluon vertex is a feature of QCD, with no parallel in QED. All partons are assumed massless. Note that the gluon structure function appears on both sides of the equation. There are similar equations for the quark and antiquark structure functions, given by Peskin and Schroeder, but since we really won't do anything with them, we will not list them here. It is useful, however, to know that such things exist. For completeness, we list the two splitting functions:

$$P_{g\leftarrow q}(z) = \frac{4}{3}\left\{\frac{1+(1-z)^2}{z}\right\}. \tag{5.97}$$

Apart from the color factor 4/3, this equation is the same as the splitting in QED $e^- \to e^- + \gamma$, as is shown by Peskin and Schroeder. The gluon splitting function, on the other hand, is unique to QCD, and more complicated:

$$P_{g\leftarrow g}(z) = 6\left\{\frac{1-z}{z} + \frac{z}{(1-z)_+} + z(1-z) + \left(\frac{11}{12} - \frac{n_f}{18}\right)\delta(1-z)\right\}. \tag{5.98}$$

The '+' on the second term is a prescription for handling the singularity at $z=1$, namely

$$\int_0^1 \frac{f(z)}{(1-z)_+} = \int_0^1 \frac{f(z)-f(1)}{1-z}, \tag{5.99}$$

where $f(z)$ is well behaved in the region $0 \le z \le 1$.

Fits to the experimental data like the ones shown in Figure 5.5 are consistent with the Altarelli–Parisi equations. The A–P equations account for the scaling violation — the Q^2 dependence of the quark structure functions — but do not predict the x dependence of $f_f(x,Q)$. The x dependence of the parton distribution functions is non-perturbative QCD, as indicated by the simple model in terms of quark binding energy distributions discussed earlier in this chapter. This feature of hadron structure is a bit awkward, in that we know corrections to the bound state, but we cannot describe that state at zero order. The presence of Λ in α_s above makes this fairly obvious, since the pion mass is less than 200 MeV. An imperfect analogy would be the non-relativistic hydrogen atom, for which we could calculate the relativistic corrections for the fine structure, but we did not know the energy levels of the Bohr formula. CTEQ and other groups that derive the parton distribution functions from experimental data assume some simple polynomial form for $f_i(x)$, and then fit the coefficients in the polynomial. This gives good fits to the experimental data, as shown in Figure 5.6, but that does not mean that the forms are 'correct'. An example for the valence up quarks in the proton, calculated for $Q^2 = 5\,\mathrm{GeV}^2$ [10], is

$$xu_v(x) = 1.8x^{0.5}(1 - x^{1.5})^{3.5}. \tag{5.100}$$

The normalization constant comes from $\int_0^1 u_v(x)dx = 2$, and as Equation (5.100) shows, u(0) diverges, so some care is needed to define the integral. The fraction of the proton momentum carried by the valence up quarks in this fit is about 0.16, and the distribution $xu_v(x)$ peaks around $x = 0.22$ with a height of 0.6, in agreement with Figure 5.5, that shows the sum $xu_v(x) + x\bar{u}(x) = 0.7$ around $x = 0.2$.

Equation (5.100) is an early version of $xu_v(x)$. Should you wish to calculate some cross-section in the parton model, use the most up-to-date forms given by CTEQ. Although the reader may not be convinced from the description of the parton model offered here, the model actually works as a practical method for calculating high energy collision phenomena. Very large Monte Carlo computer libraries, like PYTHIA, downloadable from the Lund University website (home.thep.lu.se), calculate just about any reaction, like $p + \bar{p} \to W^{\pm} +$ Anything using the parton model. Parton distribution functions are contained in the library. If you give it a luminosity, it will return an event rate that is in agreement with the experiment. PYTHIA is widely used to design experiments, estimate yields and backgrounds, and compare predictions with experimental results. The program is well documented and user friendly, and can be executed standalone, all by itself. Download it and give it a try!

5.7. Problems

5.1 Verify Equations (5.10) and (5.11).

5.2 Using the elastic scattering relation $-\frac{q^2}{2m\nu} = 1$, confirm the kinematic relation Equation (5.14):

$$\frac{E'}{E} = \frac{1}{1 + \frac{2E\sin^2(\theta/2)}{M}},$$

and obtain Equation (5.15) from Equation (5.12).

5.3 Derive the Mott cross-section Equation (5.17). Suggested procedure: put the electron mass back in the electron tensor $E^{\mu\lambda}$, and contract it with a tensor $M_{\mu\lambda}$ for which $p = p' = (M, 0, 0, 0)$. This should give $E^{\mu\lambda}M_{\mu\lambda} = 32M^2(E^2 + m_e^2 + k^2\cos\theta)$. Obtain $\overline{|\langle f|\,\mathrm{T}\,|\mathrm{i}\rangle\,|^2}$, averaging over initial spins (our heavy target has spin 1/2). Now modify Equation (5.12) for one outgoing particle in the final state, the target still being at rest as follows:

$$d\sigma = \frac{2\pi\delta(E - E')}{2M}\frac{d^3k'}{(2\pi)^3 2E'}\overline{|\langle f|\,\mathrm{T}\,|\mathrm{i}\rangle\,|^2}\frac{1}{4kM}, \tag{5.101}$$

where $1/2M$ normalizes the spinor, the last factor is the flux, and $k = k'$ is the magnitude of the electron momentum. In the matrix element $E^2(1 + \frac{m_e^2}{E^2} + v^2\cos\theta) = 2E^2(1 - v^2\sin^2(\theta/2))$; so after some cancellation and integrating over the δ function, one obtains

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2 E^2}{4k^4\sin^4(\theta/2)}(1 - v^2\sin^2(\theta/2)). \tag{5.102}$$

As mentioned in the text, in the limit $v \to 0$, this is Rutherford's formula for scattering of α particles from gold, after multiplying by $4Z^2$, and for gold $Z = 79$. It is interesting that we started out with two spin 1/2 Dirac particles, with charges and magnetic moments, scattering via electric and magnetic forces, but the formula we wound up with in the non-relativistic limit is the same as for spinless α particles. In fact it is the same as Rutherford derived using classical mechanics. In the Mott form, the factor $(1 - v^2 \sin^2(\theta/2))$ is a helicity rule for the electron spin, that becomes one in the NR limit.

5.4 Evaluate the traces in Equation (5.24) to obtain $M^{\mu\lambda}$ in Equation (5.25). Then contract $M^{\mu\lambda}$ with the electron tensor $E_{\mu\lambda}$ to obtain Equation (5.27).

5.5 Check the Rosenbluth formula for the $(\kappa F_2)^2$ term. The part of $M^{\mu\lambda}$ of interest is obtained by setting $F_1 = 0$. Then

$$\begin{aligned} M^{\mu\lambda} = 4(\kappa F_2)^2 \Big\{ & p'^{\mu} p^{\lambda} + p'^{\lambda} p^{\mu} + (M^2 - p \cdot p') g^{\mu\lambda} \\ & + \left(\frac{p \cdot p'}{4M^2} - \frac{3}{4} \right) (p + p')^{\mu} (p + p')^{\lambda} \Big\}. \end{aligned} \tag{5.103}$$

Contracting with the electron tensor and dropping the factor $(\kappa F_2)^2$ gives

$$\begin{aligned} E_{\mu\lambda} M^{\mu\lambda} = 32 \Big\{ & k \cdot p' k' \cdot p + k \cdot p k' \cdot p' - M^2 k \cdot k' \\ & + \left(\frac{p \cdot p'}{4M^2} - \frac{3}{4} \right) (k \cdot (p + p') k' \cdot (p + p') \\ & - k \cdot k' (p + p')^2 / 2) \Big\}. \end{aligned} \tag{5.104}$$

According to Equation (5.29), this expression should equal

$$\begin{aligned} & 64 M^2 E E' \left(-\frac{q^2}{4M^2} \cos^2(\theta/2) - \frac{q^2}{2M^2} \sin^2(\theta/2) \right) \\ & \quad = 4q^2 (q^2 - 4EE'). \end{aligned} \tag{5.105}$$

This exercise is to show that this is indeed true. One more intermediate step to check your algebra, and you should be able to take it from here:

$$E_{\mu\lambda} M^{\mu\lambda} = 32 \left\{ \frac{M^2 (E - E')^2}{2} - \frac{q^2 (E + E')^2}{8} + \frac{q^6}{32 M^2} \right\}. \tag{5.106}$$

There are many equivalent ways to express the invariant forms obtained from the trace theorems. We have shown that

Rosenbluth's compact formula is one of them. But we have to admire Rosenbluth's derivation in the first place.

5.6 Show that Equation (5.43) satisfies Equation (5.42).

5.7 Verity Equation (5.44) by contracting $W^{\mu\lambda}$ with the electron tensor, and then use the laboratory frame kinematics to obtain Equation (5.45).

5.8 Verify the virtual photon polarization sum rule Equation (5.58), and that the $q^\mu q^\nu$ term does not couple to the vector current $\bar{u}(k')\gamma^\lambda u(k)$.

5.9 Verify the decomposition of the inelastic electron–proton cross-section in terms of the cross-sections for transverse and longitudinal photons, Equations (5.65) and (5.66).

5.10 Obtain Equation (5.80) from Equation (5.79).

5.11 Feynman's definition $x = p_i/P$ is given in terms of invariant variables as $x = -t/(s+u)$. So, $F_2(x)$ in Equation (5.85) is directly measurable. Calculate the numerical factor in nanobarns multiplying $F_2(x)$ for 20 GeV electrons incident on a hydrogen target, and scattered electrons at $\theta = 10°$ and $E' = 10\,\mathrm{GeV}$. Calculate x. Calculate the scattered electron current multiplier $I_s F_2(x) dx dy$ for $100\,\mu$ amp incident beam, and a 50 cm long liquid hydrogen target. (I get $I_s = 36$ picoamps.) Measure the scattered current, know your acceptance widths $dxdy$, and you have $F_2(x)$.

5.12 Plot the parton distribution functions for the proton given by Eichten *et al.* [10].

$$xu_v(x) = 1.8x^{0.5}(1 - x^{1.5})^{3.5},$$

$$xd_v(x) = 0.67x^{0.4}(1 - x^{1.5})^{4.5},$$

$$x\bar{u}(x) = x\bar{d}(x) = 0.18(1 - x)^{8.5},$$

$$xs(x) = 0.08(1 - x)^{8.5},$$

$$xg(x) = (2.62 + 9.17x)(1 - x)^{5.9}.$$

The valence quark distributions vanish at $x = 0$, but the others do not. Each of the parton distribution functions, the above formulas divided by x, is singular at $x = 0$. The fits were made for $Q^2 = 5$ GeV^2 and $\Lambda = 200\,\mathrm{MeV}$ in the formula for α_s (Equation (5.95)). The Altarelli–Parisi equations were used. Note that the strange quarks are smaller than the up or down quarks in the sea by about a factor or two.

References

[1] M. N. Rosenbluth, *Phys. Rev.* **79**, 615 (1950).
[2] F. Halzen and A. Martin, *Quarks and Leptons* (John Wiley and Sons, New York, 1984).
[3] C. F. Perdrisat, V. Punjabi, and M. Vanderhaegen, Nucleon electromagnetic form factors, (arXiv:hep-ph/0612014v2, 12 September 2007).
[4] R. C. Walker *et al.*, *Phys. Rev. D* **49**, 5671 (1994).
[5] J. D. Bjorken, *Phys. Rev.* **179**, 1547 (1969).
[6] J. P. Berge *et al.*, *Zeit für Physik C* **49**, 187 (1991).
[7] S. Drell and T.-M. Yan, *Phys. Rev. Lett.* **25**, 316 (1970).
[8] J. G. Branson *et al.*, *Phys. Rev. Lett.* **38**, 1334 (1977).
[9] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, New York, 1995).
[10] E. Eichten, I. Hinchliffe, K. Lane, and C. Quigg, *Rev. Mod. Phys.* **56**, 579 (1984).

This page intentionally left blank

Chapter 6

Weak Interactions

6.1. Nuclear Beta Decays

When Becquerel exposed his photographic plates to pitchblende, the exposure was by γ rays from nuclear de-exitation and electrons from beta decay, all coming from the daughter nuclei in the uranium decay chain. All nuclei heavier than lead decay by emitting α particles, but a 5-MeV α particle has a range in air of about 10 cm, so it would not get through the paper covering Becquerel's plate. Half lives of heavy nuclei for α decay range from microseconds to millions of years, but are not caused by the weak interaction. The α particle bounces around in the nucleus, waiting for the chance to tunnel through the Coulomb barrier. The explanation of the dependence of α particle lifetimes on kinetic energy and daughter nucleus atomic number given by Gamow based on non-relativistic wave mechanics was one of the early indications that quantum mechanics could be applied to the nucleus.

Nuclear beta decays are different. Electron energies are typically a few MeV, like the α particles, but there was clear evidence from the uncertainty principle that the electron was not in the nucleus waiting to get out. Using $\Delta x \Delta p \sim \hbar$, if $\Delta p = 2$ MeV/c, then $\Delta x = 10^{-11}$ cm, much larger than a nuclear radius. This is not a problem for the α particle, because $M_\alpha = 3725\,\mathrm{MeV}$, so a 5-MeV α particle has a momentum of 200 MeV/c, and that fits inside a nucleus just fine; α particles from a radioactive source can determine orbits in dipole magnets used to bend 200 MeV/c pions, taking $Z = 2$ into account, of course.

If the electron was not hanging around in the nucleus before emission, it had to be created at the instant of beta decay, in analogy with the creation of a photon in an atomic transition. Creation of particles requires field theory. So, we anticipate that any explanation of beta decay has to involve field operators. A glance at any table of isotopes shows that all of the atomic numbers from $Z = 0$ (the neutron) to $Z = 94$ (plutonium) beta decay, with the exception of Ytterbium, $Z = 70$, that has no radioactive isotopes! Leaving Ytterbium aside, it is clear that there is no lack of sources of information on nuclear beta decays. Unlike electrodynamics, however, the weak interaction is a phenomenon without any classical analog, so we have only quantum theory to guide us in constructing an explanation.

The neutrino hypothesis explained the continuous energy distribution of electrons. For a decay $M(Z, N) \to M(Z+1, N-1)$, changing a neutron into a proton inside the nucleus, the electron kinetic energy ranges from zero to $\Delta M = M(Z, N) - M(Z+1, N-1)$. Here, we are using atomic masses, that are listed in tables of isotopes. The atomic masses are given in atomic mass units (makes sense), and one amu = 931.5 MeV. Electron binding energies can be neglected, but not electron masses. However, in an electron beta decay, the extra electron mass automatically appears in $M(Z + 1, N - 1)$, so the maximum electron kinetic energy is indeed ΔM. This formula has to be adjusted for positron emission: max positron kinetic energy $= M(Z, N) - M(Z - 1, N + 1) - 2m_e$.

Equation (4.13) for the decay rate is reproduced here for convenience for a three-particle final state:

$$\begin{aligned} d\Gamma_{i\to f} = {} & \frac{(2\pi)^4}{2m}\delta(m - E_1 - E_2 - E_3)\delta(\vec{p_1} + \vec{p_2} + \vec{p_3})|\langle f|\,T\,|i\rangle|^2 \\ & \times \frac{d^3p_1}{(2\pi)^3 2E_1}\frac{d^3p_2}{(2\pi)^3 2E_2}\frac{d^3p_3}{(2\pi)^3 2E_3}, \end{aligned} \tag{6.1}$$

where m is the mass of the parent nucleus, and $E_3 = m_3$ is the mass of the daughter nucleus (not the atomic masses). Even for the decay of the neutron, the nucleus in the final state has infinite mass compared to the kinetic energy released, so that momentum can be conserved with zero recoil energy. Integration over d^3p_3 then gives:

$$d\Gamma_{i\to f} = \frac{(2\pi)^4}{2m}\delta(m - m_3 - E_1 - E_2)|\left\langle f\right| T \left|i\right\rangle|^2 \times \frac{d^3p_1}{(2\pi)^3 2E_1}\frac{d^3p_2}{(2\pi)^3 2E_2}\frac{1}{(2\pi)^3 2m_3}. \tag{6.2}$$

The T matrix element for neutron decay we write as follows:

$$\left\langle pe^-\bar{\nu}_e\right| T \left|n\right\rangle = \frac{G_F\cos(\theta_c)}{\sqrt{2}}(\bar{u}_p\gamma^\mu(1-\gamma^5)u_n) \times (\bar{u}_e\gamma_\mu(1-\gamma^5)v_{\nu_e}), \tag{6.3}$$

where G_F is the Fermi coupling constant, and θ_c is the Cabibbo angle introduced in Chapter 1, both to be determined by experiment, the factor $1/\sqrt{2}$ is traditional, and the neutron and proton are treated as Dirac spin 1/2 particles. That part of the equation is problematic, since a Dirac spin 1/2 particle has a magnetic moment proportional to its charge with g factor 2. That would mean that $\mu_n = 0$, and $\mu_p = +e/2m_p$. Neither one of these relations is correct. The neutron moment is $\mu_n = -1.9(e/2m_p)$, and the proton moment is $\mu_p = +2.8(e/2m_p)$, both indicative of substructure, and explained by the quark model, as discussed in Chapter 1. All of the spin 1/2 neutral baryons have negative magnetic moments.

The charge of the left-hand baryon current shown in Figure 6.1 changes sign from zero to +1, and the right-hand lepton current also changes sign from zero to −1. In this formula, there is no intermediate exchanged particle. The interaction occurs at one point. But the reaction would be consistent with the exchange of a very heavy negatively charged intermediate boson, and this is the content of the electroweak theory to be discussed in detail in the next two chapters. The nucleons are replaced by quarks. This process is also shown in Figure 6.1. In the new notation, the beta decay of the neutron would be written:

$$\left\langle ue^-\bar{\nu}_e\right| T \left|d\right\rangle = \frac{g^2\cos(\theta_c)}{2}\bar{u}_{uL}\gamma^\mu u_{dL} \times \frac{i(-g_{\mu\nu} + (q_\mu q_\nu)/(M_W)^2)}{q^2 - M_W^2}\bar{u}_{eL}\gamma^\nu v_{\nu_e R}. \tag{6.4}$$

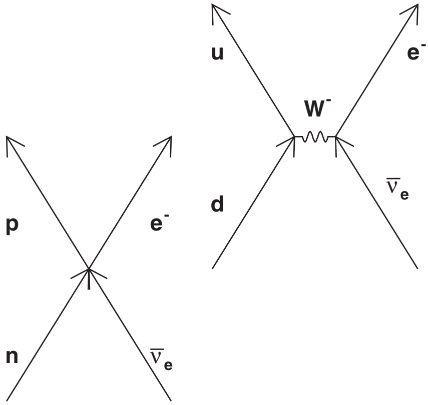


Figure 6.1: Diagrams for neutron decay. On the left side, a point interaction changes a neutron into a proton, and a neutrino into an electron. The neutrino is shown as an antineutrino in the final state. On the right side is the electroweak theory picture of neutron decay, where a down quark becomes an up quark, emitting a W^- intermediate boson, that then decays into the $e^-\bar{\nu}_e$ pair in the final state.

The coupling constant g is related to G_F by

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}. \tag{6.5}$$

The factors $(1-\gamma^5)$ are replaced by left-handed spinors and right-handed antispinors: $((1-\gamma^5)/2)u(p) = u_L(p)$; $((1-\gamma^5)/2)v(p) = v_R(p)$ for massless spin 1/2 particles. These projection operators have been mentioned in Section 4.10, and we will demonstrate this below. The W boson propagator for the momentum transfer $q^2 \ll M_W^2$ reduces to $ig_{\mu\nu}/M_W^2$. The factors of two from the spin projection operators plus the factor of 2 in the denominator of Equation (6.4) combine for the factor of 8 in Equation (6.5). The electroweak theory relates g to the elementary charge:

$$g\sin(\theta_W) = e. \tag{6.6}$$

Here, θ_W, called the Weinberg angle after Steven Weinberg, is a mixing parameter of the theory. Note that while θ_W is not known, and must be determined by experiment, $|\sin\theta_W| \leq 1$, so the relationship between the weak and electromagnetic interactions is not completely arbitrary. The Weinberg angle can be measured in many ways, as it occurs throughout electroweak theory. One method, to be discussed in Chapter 7, is the weak neutral current couplings to leptons and quarks. The world average accepted values for θ_W and M_W are

$$\sin^2(\theta_W) = 0.23122(4); M_W = 80.379(12)\,\text{GeV}/c^2. \tag{6.7}$$

Substituting these numbers together with the fine structure constant $\alpha = 1/137.036$ into Equation (6.5) gives a value for the Fermi coupling constant:

$$G_F = 1.085 \times 10^{-5}\,\text{GeV}^{-2}. \tag{6.8}$$

The Particle Data Group quotes a world average number $G_F = 1.166 \times 10^{-5}\,\text{GeV}^{-2}$, about 8% larger. The discrepancy can be explained by electroweak radiative corrections that are generally several percent, compared to less than 1% obtained in QED. G_F can also be determined from the neutron lifetime, the lifetimes of other particularly simple beta decays, like spin 0 → 0 with no parity change nuclear transitions, or the pion beta decay where the pions are also a 0 → 0 transition, $\pi^+ \rightarrow \pi^0 e^+ \nu_e$, and from the lifetime of the muon which we shall work out presently. For the neutron lifetime one must know the Cabibbo mixing angle, quoted in Chapter 1 as 13°, giving $\cos(\theta_c) = 0.974$.

Let us return to Equation (6.3), where neutron beta decay is described as a four Fermion interaction among Dirac particles. The initial state and final state nucleons are both at rest. It is useful to reproduce here some of the formulas of Dirac theory already derived. The spinors for particles are

$$u^{(1)}(p) = \sqrt{E+m}\begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \end{pmatrix}, \quad u^{(2)}(p) = \sqrt{E+m}\begin{pmatrix} 0 \\ 1 \\ \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \end{pmatrix}, \tag{6.9}$$

and those for antiparticles are

$$v^{(1)}(p) = \sqrt{E+m}\begin{pmatrix} \frac{p_x - ip_y}{E+m} \\ \frac{-p_z}{E+m} \\ 0 \\ 1 \end{pmatrix}, \quad v^{(2)}(p) = \sqrt{E+m}\begin{pmatrix} \frac{p_z}{E+m} \\ \frac{p_x + ip_y}{E+m} \\ 1 \\ 0 \end{pmatrix}. \tag{6.10}$$

The neutron and proton are both particles at rest, so the bottom two components vanish. The Dirac operators in our representation in 2×2 form (each entry is a 2×2 matrix) are

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ \vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}, \ \gamma^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \gamma^0\gamma^5 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{6.11}$$

$$\vec{\gamma}\gamma^5 = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix}, \ \frac{1-\gamma^5}{2} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \ \frac{1+\gamma^5}{2} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{6.12}$$

It is clear from these relations that only γ^0 and $\vec{\gamma}\gamma^5$ contribute to the nucleon matrix element, they being the only operators that have an entry in the first box.

For the leptons it is safe to set the antineutrino mass to zero. So, if we choose the axis of quantization of the antineutrino spin, the z axis, to coincide with the momentum direction, then the antineutrino spinors are

$$v^{(1)}(p) = \sqrt{E}\begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \quad v^{(2)}(p) = \sqrt{E}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \tag{6.13}$$

In the electron high energy limit (which is not always a good approximation for nuclear beta decays), with the same choice of momentum direction the spinors become:

$$u^{(1)}(p) = \sqrt{E}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad u^{(2)}(p) = \sqrt{E}\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}. \tag{6.14}$$

The following are relations for the helicity projection operators that follow from the definitions:

$$\begin{aligned}
&\frac{1+\gamma^5}{2}+\frac{1-\gamma^5}{2}=1, \quad (1+\gamma^5)\times(1-\gamma^5)=0,\\
&\frac{1\pm\gamma^5}{2}\times\frac{1\pm\gamma^5}{2}=\frac{1\pm\gamma^5}{2},
\end{aligned} \tag{6.15}$$

and

$$\begin{aligned}
(1+\gamma^5)v^{(1)}(p)&=(1+\gamma^5)u^{(2)}(p)=(1-\gamma^5)v^{(2)}(p)\\
&=(1-\gamma^5)u^{(1)}(p)=0,
\end{aligned} \tag{6.16}$$

and

$$\begin{aligned}
&\frac{1+\gamma^5}{2}u^{(1)}(p)=u^{(1)}(p), \quad \frac{1+\gamma^5}{2}v^{(2)}(p)=v^{(2)}(p),\\
&\frac{1-\gamma^5}{2}u^{(2)}(p)=u^{(2)}(p), \quad \frac{1-\gamma^5}{2}v^{(1)}(p)=v^{(1)}(p).
\end{aligned} \tag{6.17}$$

So, the operators $(1\pm\gamma^5)/2$ project out positive and negative helicities for relativistic electrons, and the opposite for relativistic positrons. The action of the projection operator for $v<1$ can be obtained from the expectation value:

$$\left\langle u^{\dagger(2)}(p)\right|\frac{(1-\gamma^5)}{2}\left|u^{(2)}(p)\right\rangle=\frac{E+p}{2E}=\frac{1}{2}(1+v). \tag{6.18}$$

So, for a slow electron, the helicity is multiplied by v. This will be important later on, when we are looking at decays like $\pi^+\to\mu^++\nu_\mu$, that violate the helicity rules. Neutrinos and antineutrinos are always relativistic. The choice of $(1-\gamma^5)$ in the weak interaction T matrix is consistent with experiment. Left-handed leptons and right-handed antileptons are preferred by the weak interaction. This preference is explicitly parity violating, since spin is an axial vector and momentum is a polar vector, the helicity defined as $\vec{\sigma}\cdot\hat{p}$ is a pseudo scalar, that changes sign under space reflection. The general form of the parity violating weak interaction is called 'Vector minus Axial Vector', or '$V-A$' for short. In nuclear beta decay terminology the vector part is called 'Fermi', and the axial vector part

'Gamow–Teller'. Each term separately would be parity conserving, like electrodynamics.

Equation (6.3) when multiplied contains the sum of four terms, that we call A, B, C, and D, all multiplied by $G_F \cos(\theta_c)/\sqrt{2}$:

$$\langle pe^- \bar{\nu}_e | T | n \rangle = \frac{G_F \cos\theta_c}{\sqrt{2}} (\mathrm{A} + \mathrm{B} + \mathrm{C} + \mathrm{D}), \tag{6.19}$$

where

$$\mathrm{A} = (\bar{u}_p \gamma^\mu u_n)(\bar{u}_e \gamma_\mu v_\nu), \quad \mathrm{B} = -(\bar{u}_p \gamma^\mu u_n)(\bar{u}_e \gamma_\mu \gamma^5 v_\nu), \tag{6.20}$$

and

$$\mathrm{C} = -\lambda(\bar{u}_p \gamma^\mu \gamma^5 u_n)(\bar{u}_e \gamma_\mu v_\nu), \quad \mathrm{D} = \lambda(\bar{u}_p \gamma^\mu \gamma^5 u_n)(\bar{u}_e \gamma_\mu \gamma^5 v_\nu). \tag{6.21}$$

An unknown parameter λ has been inserted multiplying the axial vector part of the nucleon current, allowing for effects caused by the fact that the nucleons are not Dirac particles. We shall see that the electron–antineutrino angular correlation depends upon λ, and a recent measurement is $\lambda = 1.284 \pm 0.014$ [1]. The antineutrino is not observed; the angular correlation is inferred from the angle between the electron and the recoil proton, that can be measured in a slow neutron beam.

The sum over spins of the square of the T matrix element is

$$\sum_{\text{spins}} |\langle pe^- \bar{\nu} | T | n \rangle|^2 = \frac{G_F^2 \cos^2(\theta_c)}{2} \sum_{\text{spins}} (\mathrm{A}^\dagger + \mathrm{B}^\dagger + \mathrm{C}^\dagger + \mathrm{D}^\dagger) \times (\mathrm{A} + \mathrm{B} + \mathrm{C} + \mathrm{D}). \tag{6.22}$$

This expression contains a horrifying 16 terms! Fortunately, only four of them for neutron decay have non-zero trace. All terms in the baryon current with an odd number of γ^5's vanish (two γ^0's), and those for the lepton current cancel in pairs, because the $\mathrm{Tr}(\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \gamma^5) = 4i\epsilon^{\mu\nu\rho\sigma}$, Equation (4.48), and the final expression is real. If there are an even number of γ^5's, they can be multiplied out, so that the traces are completely free of them. The results are as follows:

$$\mathrm{Tr}(\mathrm{A}^\dagger \mathrm{A}) = \mathrm{Tr}(\mathrm{B}^\dagger \mathrm{B}) = 32m_p^2(2E_eE_\nu - p_e \cdot p_\nu), \tag{6.23}$$

and

$$\mathrm{Tr}(\mathrm{C}^\dagger \mathrm{C}) = \mathrm{Tr}(\mathrm{D}^\dagger \mathrm{D}) = 32m_p^2\lambda^2(2E_eE_\nu + p_e \cdot p_\nu). \tag{6.24}$$

We have set $m_p = m_n$, which is OK to about two parts in a thousand. The angular correlation is contained in the four vector product: $p_e \cdot p_\nu = E_eE_\nu - \vec{p}_e \cdot \vec{p}_\nu$, and clearly vanishes if $\lambda = 1$. Assembling all of the parts gives:

$$\sum_{\text{spins}} |\left\langle pe^-\bar{\nu}_e\right| T \left|n\right\rangle|^2 = \frac{G_F^2 \cos^2(\theta_c)}{2} 64m_p^2E_eE_\nu(1+3\lambda^2) \times \left(1 + \frac{1-\lambda^2}{1+3\lambda^2}|v_e|\cos(\theta_{e\nu})\right), \tag{6.25}$$

where the electron velocity $|v_e| = |\vec{p}_e|/E_e$ multiplies the angular correlation. Equation (6.25) is the $V-A$ theory prediction for the decay of unpolarized neutrons. For polarized neutrons there is a correlation term $\vec{\sigma}_n \cdot \vec{p}_e/E_e$. For real λ the coefficient of this spin dependence is $A = -2(\lambda^2 + \lambda)/(1 + 3\lambda^2)$. In order to relate theory to the neutron lifetime we must evaluate the phase space integral multiplied by E_eE_ν. The factors $2E_eE_\nu$ and $2m_p^2$ come from our normalization of the Dirac spinors (Equation (6.9)), divided in each case by two to make up for the missing factor in the helicity projection operator. The angular correlation term is linear in $\cos(\theta_{e\nu})$ and vanishes when integrated over the solid angle. Before we do that, however, we should make a couple of points that can be directly obtained from the work we have already done.

In nuclear beta decays either a neutron inside a nucleus becomes a proton, emitting an electron and an antineutrino, or a proton becomes a neutron (not energetically allowed for free protons), emitting a positron and a neutrino. Light nuclei are often proton rich, and are neutrino emitters. The fuel cycle in the sun is such a process; the sun emits neutrinos that can make electrons in matter by the reaction $\nu_e + n \to e^- + p$. Fission products on the other hand are often neutron rich, as the fissioning heavy elements have too many neutrons to begin with, so nuclear power reactors are sources of antineutrinos, that make positrons in matter by the reaction $\bar{\nu}_e + p \to e^+ + n$. Nuclear beta decays are complicated by nuclear structure considerations. One proton becomes one neutron, or vice versa, and the daughter nucleus has to accommodate the altered nucleon. The term $\mathrm{A}^\dagger\mathrm{A}$ in Equation (6.23) is pure Fermi, or vector, and for a complex nucleus the spin-parity selection rules are $\Delta J = 0$, no parity change. Similarly the $\mathrm{D}^\dagger\mathrm{D}$ term in Equation (5.24) is pure Gamow–Teller, or axial vector, for which the nuclear selection rules are $\Delta J = \pm 1, 0$, not $0 \to 0$, again with no parity change.

There are several light nuclei that are $0^+ \to 0^+$ beta emitters, and therefore examples of pure vector decay. One interesting example is $^{14}\mathrm{O} \to {}^{14}\mathrm{N}^* + e^+ + \nu_e$. The ground state of ^{14}O decays to the first excited state of ^{14}N. Both states have spin/parity 0^+. The ground state of ^{14}C, one atomic number below nitrogen, also has spin/parity 0^+. The ground state of ^{14}N has spin one. Alternating intensities in the rotational Raman spectrum of diatomic nitrogen showed that the spin of the nucleus was one, a result that at the time eliminated the possibility that ^{14}N was composed of 14 protons and seven electrons, which would lead to 1/2 integer spin. ^{14}N is a stable odd–odd nucleus, with seven protons and seven neutrons. Other stable odd–odd light nuclei are ^{2}H spin 1, ^{6}Li spin 1, and ^{10}B spin 3. A reasonable model for ^{14}N could be a deuteron outside the spin zero ^{12}C core, and for the first excited state the deuteron is broken up by flipping a spin (the binding energy is about 2.2 MeV) to give an np state with spin 0, that can have I-spin one. In this picture, the ground state of ^{14}C with two neutrons outside the core, $^{14}\mathrm{N}^*$ with np outside the core, and the ground state of ^{14}O with two protons outside the core form an I-spin triplet, and one of these two outer protons beta decays. The nuclear matrix element would then be replaced by the I-spin lowering operator, namely $\sqrt{2}$, assuming that the wave functions overlap each other 100%. The relevant energies are given Table 6.1.

Hence the end point for the positron decay of ^{14}O is $KE_{\max} = 5.14 - 1.022 - 2.31 = 1.81\,\mathrm{MeV}$. The last three levels, the I-spin

Table 6.1: Atomic masses for the $I = 1$ triplet relative to the ^{14}N ground state.

Atom	Mass (MeV)	Spin/parity	I-Spin	I_z
^{14}N	0	1^+	0	0
^{14}N*	2.31	0^+	1	0
^{14}C	0.156	0^+	1	-1
^{14}O	5.14	0^+	1	$+1$

triplet, would have the same energy if I-spin were an exact symmetry. There is about a 2 MeV wobble about the mean. Based on this argument we expect the square of the T matrix element for the positron decay of ^{14}O to be:

$$|\left\langle ^{14}\mathrm{N}^*\mathrm{e}^+\nu_\mathrm{e}\right| T \left|^{14}\mathrm{O}\right\rangle|^2 = \frac{G_F^2 \cos\theta_c^2}{2} 64 m_p^2 (2E_e E_\nu - p_e \cdot p_\nu). \quad (6.26)$$

We now turn to integrating the phase space term, keeping in mind that a proton has decayed into a neutron. The fact that each one is in a nucleus is just taken care of by the $\sqrt{2}$. So we may write the differential decay rate:

$$d\Gamma_{i\to f} = \frac{2G_F^2\cos^2\theta_c}{(2\pi)^5}\delta(Q - E_e - E_\nu)d^3p_e d^3p_\nu(1 + |v_e|\cos\theta_{e\nu}). \quad (6.27)$$

The argument of the δ function is the total kinetic energy plus the electron mass: $Q = 2.31\,\mathrm{MeV}$. First integrate over the positron–neutrino solid angle, giving 4π and killing the angular correlation term, and then integrate over the positron solid angle, giving another factor of 4π. Integrate $\int E_\nu^2 dE_\nu \delta(Q - E_e - E_\nu)$. Now define $x = E_e/Q$, the range of x is $m_e/Q \leq x \leq 1$. We now have the integral:

$$\Gamma_{i\to f} = \frac{G_F^2\cos^2\theta_c}{\pi^3} Q^5 \int_a^1 x(1-x)^2\sqrt{x^2 - a^2}dx, \quad (6.28)$$

where $a = m_e/Q = 0.22$. The integral can be expanded into standard forms. The result is

$$\int_a^1 x(1-x)^2\sqrt{x^2-a^2}dx = \sqrt{1-a^2}\left(\frac{1}{30}-\frac{3a^2}{20}-\frac{2a^4}{15}\right) + \frac{a^4}{4}\ln\left(\frac{1+\sqrt{1-a^2}}{a}\right) = 0.026. \tag{6.29}$$

The square of the Fermi coupling constant (without some small corrections) in MeV^{-4} is given in terms of the lifetime $\tau = 102\,\text{s}$ of ^{14}O:

$$G_F^2 = \frac{\pi^3\hbar}{0.026\cos^2(\theta_c)Q^5 \times \tau}. \tag{6.30}$$

Substituting in the numbers ($\hbar = 6.58 \times 10^{-22}$MeV-s) gives:

$$G_F = 1.109 \times 10^{-11}\text{MeV}^{-2}, \tag{6.31}$$

when converted to GeV^{-2} this number is very close to the world average quoted above.

The same phase space integral can be applied to the neutron decay. $Q = 1.29\,\text{MeV}$ for neutron decay, and $a = m_e/Q = 0.4$. Equation (6.29) then becomes 0.016. The neutron lifetime has vector and axial vector terms, so Equation (6.30) is modified. We use the neutron lifetime $\tau_n = 879\,\text{s}$. The Fermi coupling constant in terms of the neutron lifetime is

$$G_F^2 = \frac{2\pi^3\hbar}{0.016\cos^2(\theta_c)Q^5(1+3\lambda^2) \times \tau_n}. \tag{6.32}$$

The extra factor of two in the numerator comes from averaging over the two spin directions of the parent neutron. These numbers give

$$G_F = 1.19 \times 10^{-11}\text{MeV}^{-2}, \tag{6.33}$$

again quite close to the world average.

There is some controversy over the exact value of the neutron lifetime. See Stephan Paul for a review [2]. The neutron lifetime is difficult to measure, and it has taken several decades of work for the number to settle down. Nearly 1000 s is just too long. Even thermal neutrons have an average speed of about 2000 m/s, so in one lifetime they travel a long way. You can of course start out with a very intense beam of thermal neutrons, and some small fraction will decay in your meter or so long detector. This is called the 'beam method'. It is also possible to store neutrons much colder than room temperature in a bottle, let them sit there for a while, and count how many are left. This is called the 'bottle method'. Both have been used, and give very accurate lifetime measurements, with errors or 1 s or so. But the two methods disagree by about 8 s (1%). The bottle method is the number quoted above, and it is the shorter of the two.

6.2. Muon Decay $\mu^- \to e^- + \nu_\mu + \bar{\nu}_e$

The T matrix element for muon decay has the same general form as neutron decay: a μ^- and a ν_μ on one leg, and an e^- and a $\bar{\nu}_e$ on the other. There is no Cabibbo angle, because the mixing applies only to the quarks, as we shall learn soon. So, we write:

$$\left\langle e^-\bar{\nu}_e\nu_\mu\right| T \left|\mu^-\right\rangle = \frac{G_F}{\sqrt{2}}\bar{u}(p_2)\gamma^\lambda(1-\gamma^5)u(p_1)\bar{u}(p_4)\gamma_\lambda(1-\gamma^5)v(p_3), \tag{6.34}$$

where we have labeled the spinors with their four momenta: $p_1 = \mu^-$, $p_2 = \nu_\mu$, $p_3 = \bar{\nu}_e$, and $p_4 = e^-$. There is a big difference in the kinematics of muon decay compared to neutron decay. All final state particles can be considered massless, and there is only one particle that is at rest, the parent muon in its rest frame. Terms involving γ^5 now contribute to the trace, so it is useful to work out some of those traces:

$$\begin{aligned}
\mathrm{Tr}(\gamma^\mu(1-\gamma^5)\not{p}_1\gamma^\nu(1-\gamma^5)\not{p}_2) &= \mathrm{Tr}(\gamma^\mu\not{p}_1\gamma^\nu\not{p}_2) - \mathrm{Tr}(\gamma^\mu\not{p}_1\gamma^\nu\gamma^5\not{p}_2)\\
&\quad - \mathrm{Tr}(\gamma^\mu\gamma^5\not{p}_1\gamma^\nu\not{p}_2) + \mathrm{Tr}(\gamma^\mu\gamma^5\not{p}_1\gamma^\nu\gamma^5\not{p}_2)\\
&= 2(\mathrm{Tr}(\gamma^\mu\gamma^\alpha\gamma^\nu\gamma^\beta) - \mathrm{Tr}(\gamma^\mu\gamma^\alpha\gamma^\nu\gamma^5\gamma^\beta))\\
&\quad \times p_{1\alpha}p_{2\beta}\\
&= 8(p_1^\mu p_2^\nu + p_1^\nu p_2^\mu - g^{\mu\nu}(p_1\cdot p_2))\\
&\quad + 8i\epsilon^{\mu\alpha\nu\beta}p_{1\alpha}p_{2\beta}.
\end{aligned} \tag{6.35}$$

We have used Equation (4.48), quoted in the text above. The muon mass appears in the completeness relation $\sum_{\text{spins}} u(p)\bar{u}(p) = \not{p} + m$, but the mass terms do not contribute to the trace. The sign has flipped in the γ^5 trace because we interchanged γ^5 and γ^β. The next handy formula involves the product of two traces of this type:

$$\begin{aligned}
&\mathrm{Tr}(\gamma^\mu(1-\gamma^5)\not{p}_1\gamma^\nu(1-\gamma^5)\not{p}_2)\mathrm{Tr}(\gamma_\mu(1-\gamma^5)\not{p}_3\gamma_\nu(1-\gamma^5)\not{p}_4)\\
&\quad = 64(p_1^\mu p_2^\nu + p_1^\nu p_2^\mu - g^{\mu\nu}(p_1\cdot p_2) + i\epsilon^{\mu\alpha\nu\beta}p_{1\alpha}p_{2\beta})(p_{3\mu}p_{4\nu}\\
&\qquad + p_{3\nu}p_{4\mu} - g_{\mu\nu}(p_3\cdot p_4) + i\epsilon_{\mu\rho\nu\sigma}p_3^\rho p_4^\sigma)\\
&\quad = 256(p_1\cdot p_3)(p_2\cdot p_4). \qquad (6.36)
\end{aligned}$$

Several terms cancel to obtain the final answer, and we have used Equation (2.30) for a two index contraction of two ϵ's:

$$-\epsilon^{\mu\alpha\nu\beta}\epsilon_{\mu\rho\nu\sigma} = -\epsilon^{\mu\nu\alpha\beta}\epsilon_{\mu\nu\rho\sigma} = 2(\delta^\alpha_\rho\delta^\beta_\sigma - \delta^\alpha_\sigma\delta^\beta_\rho). \qquad (6.37)$$

This result has the two incoming arrows on the Feynman diagram dotted into each other, and also the two outgoing arrows. Barger and Phillips [3] point out that this is a feature of $(V-A)^2$ or $(V+A)^2$, while for a mixed interaction $(V-A)(V+A)$ the momentum with incoming arrow at one W vertex is dotted into the momentum with the outgoing arrow at the other W vertex. Equation (6.36) solves our problem:

$$\sum_{\text{spins}} |\langle e^-\bar{\nu}_e\nu_\mu|\, T\, |\mu^-\rangle|^2 = \frac{G_F^2}{2}256 m_\mu E_3(p_2\cdot p_4). \qquad (6.38)$$

Particle 2 is a neutrino and 4 is the electron. We are interested in the electron energy spectrum, and that is E_4. The muon four momentum is $p_1 = (m_\mu, 0, 0, 0)$. Let us drop the index, so that m refers to the muon mass, it being the only mass in the problem. The dot product can be eliminated from this formula by using the conservation law, and the fact that the particles are massless. Thus,

$$\begin{aligned}
&p_1 = p_2 + p_3 + p_4, \quad p_2 + p_4 = p_1 - p_3,\\
&2p_2\cdot p_4 = (p_1 - p_3)^2 = m^2 - 2mE_3. \qquad (6.39)
\end{aligned}$$

So, we may write

$$\sum_{\text{spins}} |\left\langle e^- \bar{\nu}_e \nu_\mu \right| T \left| \mu^- \right\rangle = \frac{G_F^2}{2} 128 m^3 E_3 (1 - 2E_3/m). \tag{6.40}$$

The phase space, with three massless particles in the final state, is not the same as it is for neutron decay. We adopt Equation (6.1) for our present purposes:

$$\begin{aligned} d\Gamma_{i \to f} = {} & \frac{G_F^2}{16\pi^5} m^2 (1 - 2E_3/m) \delta(m - E_2 - E_3 - E_4) \delta(\vec{p}_2 + \vec{p}_3 + \vec{p}_4) \\ & \times d^3p_2/E_2 d^3p_3 d^3p_4/E_4. \end{aligned} \tag{6.41}$$

A factor of 1/2 has been added to average over muon spins. The two neutrinos are not observed, so we integrate over them. $\int d^3p_2 \delta(\vec{p}_2 + \vec{p}_3 + \vec{p}_4)$ takes care of one of the integrals, and leaving the energy δ function:

$$\begin{aligned} d\Gamma_{i \to f} = {} & \frac{G_F^2}{4\pi^4} m^2 (1 - 2E_3/m) \delta \left(m - E_3 - E_4 \right. \\ & \left. - \sqrt{E_3^2 + E_4^2 + 2E_3 E_4 \cos\theta_{34}} \right) E_4 dE_4 E_3^2 dE_3 d\Omega_{34}/E_2. \end{aligned} \tag{6.42}$$

Converting the argument of the δ function to integrate over the solid angle gives a factor of $E_2/(E_3 E_4)$:

$$d\Gamma_{i \to f} = \frac{G_F^2}{2\pi^3} m^2 E_3 (1 - 2E_3/m) dE_3 dE_4. \tag{6.43}$$

We still need to integrate over the unobserved neutrino E_3, and E_4 and E_3 are not independent variables. The maximum value of any final state energy is $E_{\max} = m/2$, and because all of the particles are massless, the maximum value of $E_3 = m/2$ independent of E_4, because the other neutrino can be parallel to the electron. However, the minimum value of E_3 cannot be zero, and balance energy and momentum, if $E_4 < m/2$. So the range of E_3 is $(m/2 - E_4) \leq E_3 \leq m/2$.

Let $y = 2E_3/m$, and $x = 2E_4/m$. Then we have:

$$\int_{m/2-E_4}^{m/2} E_3(1 - 2E_3/m)dE_3 = \frac{m^2}{4}\int_{1-x}^{1} y(1-y)dy = \frac{m^2}{4}\left(\frac{x^2}{2} - \frac{x^3}{3}\right). \tag{6.44}$$

Then

$$\frac{d\Gamma}{dx} = \frac{G_F^2 m^5}{96\pi^3} x^2(3 - 2x). \tag{6.45}$$

The integral is elementary, and the result is

$$\Gamma(\mu^- \to e^- \bar{\nu}_e \nu_\mu) = \frac{G_F^2 m^5}{192\pi^3}. \tag{6.46}$$

The Fermi coupling constant from this formula, given the muon mass $m = 105.6$ MeV, and the muon lifetime $\tau = 2.2 \times 10^{-6}$ s, is

$$G_F = 1.166 \times 10^{-11}\text{MeV}^{-2}, \tag{6.47}$$

in good agreement with the world average quoted above.

In Figure 6.2, we show a comparison of the electron energy distribution in muon decay and the positron energy distribution in ^{14}O decay. Each distribution is normalized to one. They look very different. The nuclear beta decay energy vanishes at the end points, and has a peak in the middle, while the muon decay energy rises monotonically until it hits the maximum at $x = 1$. The observed end point is not a vertical line due to experimental resolution and radiation effects, but it is true that most of the electrons have high energy.

6.3. Polarized Muons

It is interesting to work out the asymmetry distribution of electrons from polarized muons. At high electron energy, the electron is left handed, and the neutrino–antineutrino recoil spins cancel, so the electron prefers to be emitted opposite to the muon spin, as shown in Figure 6.3. The muon magnetic moment can be precessed in a perpendicular magnetic field, and the decay electrons rotate around

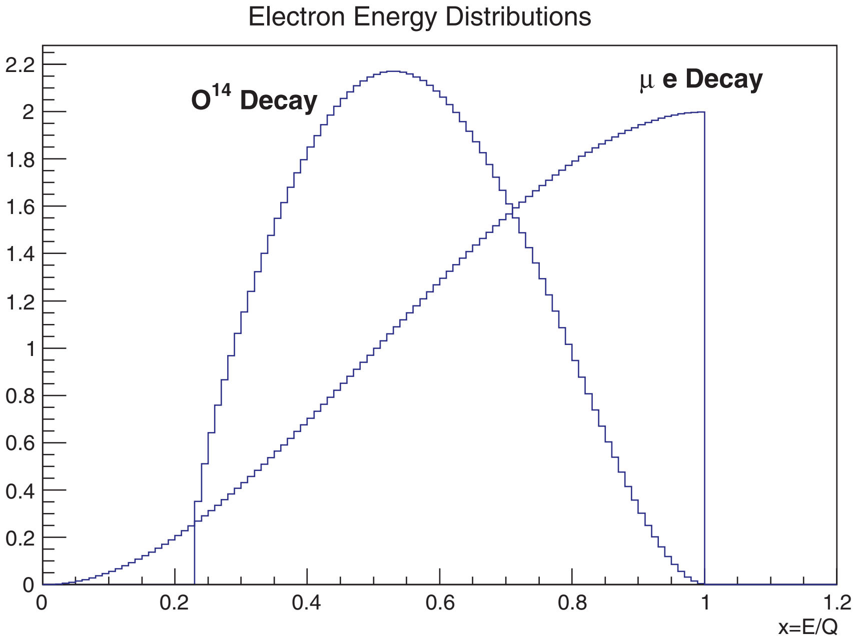


Figure 6.2: Electron spectrum from muon decay compared to positron spectrum from O^{14} decay. The end point $x = 1$ is a different energy for the two decays: 53 MeV for muon decay, and 2.31 MeV for oxygen. The oxygen spectrum is displaced from the origin by the electron mass in units of Q.

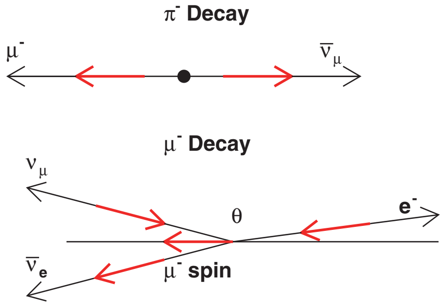


Figure 6.3: Top diagram decay of the spin zero π^- into a μ^- that is 100% polarized along its direction of motion because of the helicity of the recoil antineutrino. Spin vectors are red arrows. The bottom diagram then shows the decay of the polarized muon. The angle θ is between the muon spin direction and the electron momentum vector. The two recoil neutrino spins cancel each other.

with it. This remarkable effect was one of the first demonstrations of parity violation in the weak interactions. Polarized muons come directly from pion decay.

The electron tensor remains the same. The muon tensor gets a spin projection operator added to its completeness sum, as discussed in Chapter 4. We are therefore interested in evaluating the expression:

$$\sum_{\text{spins}} \bar{u}(p_2)\gamma^\sigma(1-\gamma^5)u(p_1)\bar{u}(p_1)\gamma^\rho(1-\gamma^5)u(p_2)$$

$$= \mathrm{Tr}\left(\not{p}_2\gamma^\sigma(1-\gamma^5)(\not{p}_1+m)\frac{(1+\lambda_1\gamma^5\not{s}_1)}{2}\gamma^\rho(1-\gamma^5)\right). \quad (6.48)$$

We retain the muon mass, and insert a spin projection operator for the muon rest frame, Equation (4.108), with a sign parameter $\lambda_1 = \pm 1$ (not the same as used in the neutron decay formula). We leave as an exercise the demonstration of this identity, which allows us to write down the answer without doing any extra work:

$$(1-\gamma^5)(\not{p}_1+m)(1+\lambda_1\gamma^5\not{s}_1)(1+\gamma^5)$$
$$= (1-\gamma^5)(\not{p}_1-\lambda_1 m\not{s}_1+m)(1+\gamma^5). \quad (6.49)$$

We may therefore make the four vector substitution $p_1 \to (p_1 - \lambda_1 m s_1)/2$ in the final formula, Equation (5.36). Thus,

$$256(p_1\cdot p_3)(p_2\cdot p_4) \to 128((p_1-\lambda_1 m s_1)\cdot p_3)(p_2\cdot p_4). \quad (6.50)$$

The unpolarized cross-section is restored by adding up two terms with $\lambda_1 = +1$ and $\lambda_1 = -1$. The author learned this trick from Barger and Phillips [4]. The same trick could be used in the electron tensor if the electron polarization in the final state were measured. The muon is at rest, so the four vector $s_1^\alpha = (0,0,0,1)$, and $s_1 \cdot p_3 = -p_{3z}$. Particle #3 is an electron antineutrino in the final state, and it escapes detection, so this angular correlation is not observed. Only the final state electron is observed, and its angular correlation with the muon spin has to be worked out from the phase space integrals.

We have already worked out the part independent of λ_1, so we have only to evaluate:

$$d\Gamma_{\text{spin}} = -\frac{128G_F^2\lambda_1}{2(2\pi)^5} m(s_1 \cdot p_3)(p_2 \cdot p_4) \left\{ \frac{d^3p_2 d^3p_3 d^3p_4}{16mE_2E_3E_4} \delta(\vec{p}_2 + \vec{p}_3 + \vec{p}_4) \right.$$
$$\left. \times \delta(m - E_2 - E_3 - E_4) \right\}. \qquad (6.51)$$

The phase space integral for this problem, with a component of one of the neutrino momenta in the matrix element, is tricky to handle. We appeal to another formula in Barger and Phillips. They define the two-neutrino invariant phase space as

$$\int d_2(PS) = \int \delta^{(4)}(X - p_2 - p_3) \frac{d^3p_2}{2p_2^0} \frac{d^3p_3}{2p_3^0}. \qquad (6.52)$$

They then write the following equation (p. 71 in *Collider Physics*):

$$\int d_2(PS) p_2^\alpha p_3^\beta = \frac{\pi}{24} \{g^{\alpha\beta} X^2 + 2X^\alpha X^\beta\}. \qquad (6.53)$$

The proof of this equation is left as an exercise, but not to worry, I will show you how to do it. The four vector X is

$$X^\alpha = p_2^\alpha + p_3^\alpha = p_1^\alpha - p_4^\alpha = (m - E_4, -p_{4x}, -p_{4y}, -p_{4z}),$$

and

$$X^2 = (p_1^\alpha - p_4^\alpha)(p_{1\alpha} - p_{4\alpha}) = m^2 \left(1 - \frac{2E_4}{m}\right). \qquad (6.54)$$

We rewrite Equation (6.51) in a form appropriate for these formulas:

$$d\Gamma_{\text{spin}} = \frac{G_F^2\lambda_1}{2\pi^5} \frac{p_{4\alpha} d^3p_4}{E_4} \left\{ \int d_2(PS) p_3^{(3)} p_2^\alpha \right\}. \qquad (6.55)$$

The sign change comes from $(s_1 \cdot p_3)$. Using Equation (6.53) gives:

$$d\Gamma_{\text{spin}} = \frac{G_F^2\lambda_1}{2\pi^5} \frac{d^3p_4}{E_4} \frac{\pi}{24} \left\{ p_4^{(3)} X^2 - 2p_4^{(3)} p_4 \cdot X \right\}, \text{ with } p_4 \cdot X = mE_4. \qquad (6.56)$$

In terms of $x = 2E_4/m$ we have after integrating over the azimuthal angle:

$$\frac{d\Gamma}{dx d\cos\theta} = -\frac{G_F^2\lambda_1}{192\pi^3}x^2(2x-1)\cos\theta. \tag{6.57}$$

Setting $\lambda_1 = +1$ for the μ^- spin along the $+z$ direction, and combining this expression with the unpolarized part, gives:

$$\frac{d\Gamma}{dx d\cos\theta} = \frac{G_F^2}{192\pi^3}x^2(3-2x)\left(1-\frac{2x-1}{3-2x}\cos\theta\right). \tag{6.58}$$

The asymmetry parameter for the electron is $\alpha = -(2x-1)/(3-2x)$, and changes sign at $x = 1/2$. For $x > 1/2$ (about 80% of the electron energy spectrum) the asymmetry parameter is negative, confirming the helicity argument given at the beginning of this calculation. For high energy negative helicity electrons, the recoil neutrino–antineutrino pair spins cancel, so the spin of the parent muon is the electron helicity, and they prefer to come out backwards. For positive muon decay, the positrons have positive helicity, so they prefer to come out forwards.

Figure 6.4 shows the data for $\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$ for polarized muons stored in a precision magnetic field. The oscillation frequency is proportional to the magnetic moment anomaly. If the muon decay asymmetry, Equation (6.58), were zero, there would be no wiggles. Equation (4.164) in mks units gives this angular frequency as

$$\omega_L - \omega_c = \frac{eBa_\mu}{m},$$

where the anomaly in terms of the g factor is

$$a_\mu = \frac{g-2}{2}. \tag{6.59}$$

Substituting in the elementary charge $e = 1.602 \times 10^{-19}\,\mathrm{C}$, $m = 1.88 \times 10^{-28}\,\mathrm{kg}$, $B = 1.45\,\mathrm{T}$, and $a_\mu = 0.001166$ gives the precession period quoted in the figure caption, $4.37\,\mu\mathrm{s}$. The drop in muon intensity over $100\,\mu\mathrm{s}$ shown in the figure is due to the time dilated muon lifetime $29.3 \times 2.2\,\mu\mathrm{s}$. Without time dilation the intensity would drop from one side of the plot to the other by 10^{-20}! This figure is an excellent demonstration of the muon lifetime, relativistic time dilation, asymmetry parameter, and anomalous magnetic moment.

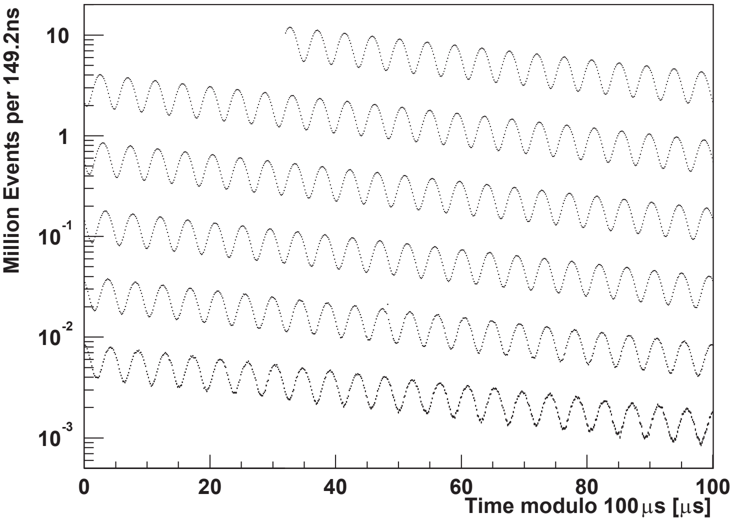


Figure 6.4: Experimental data. Polarized positive muons precessing and decaying in a 1.45 T storage ring. $E/m_\mu = 29.3$ for these muons. The laboratory lifetime $\gamma\tau_\mu = 64.4\,\mu$s, and the precession period $= 4.37\,\mu$s. Figure courtesy of Priscilla Cushman, and the Brookhaven muon $g-2$ experiment.

6.4. Pion Decay

We have learned from analysis of nuclear beta decays and muon decay that the rate scales as Q^5, where Q is the energy available to the final state leptons. This dependence is consistent with dimensional arguments, since the Fermi coupling constant $G_F \sim E^{-2}$, $G_F^2 Q^5 \sim E$, and in our units $E \sim 1/t$. This scale factor is a useful tool to estimate weak decay rates, and to discern selection rules and other effects that cause the scaling to fail. Properties of the pion triplet are summarized in Table 6.2.

A glance at the table shows that our Q^5 scaling is in trouble. The charged pion decay rate is about 100 times faster than $\mu \to e + \nu + \bar{\nu}$, yet $(m_\pi/m_\mu)^5 = 4$. Also there is the rare decay mode $\pi^+ \to e^+ + \nu_e$ that is suppressed by a factor 10^{-4}, but has the same Q value as $\pi \to \mu\nu$. The very rare decay $\pi^+ \to \pi^0 + e^+ + \nu_e$ if it were the only mode would have a lifetime of a few seconds, and in fact is consistent with ^{14}O beta decay. The lifetime of the π^0 is eight orders

Table 6.2: Properties of the pions, I-spin 1.

Symbol	Mass (MeV)	Spin/ parity	Lifetime (s)	Decay mode	$\Gamma_i/\Gamma_{\text{tot}}$	Q (MeV)
π^{+-}	139.6	0^-	2.60×10^{-8}	$\mu + \nu_\mu$	~100%	139.6
				$e + \nu_e$	1.23×10^{-4}	139.6
				$\pi^0 + e + \nu_e$	1.02×10^{-8}	4.6
π^0	134.9	0^-	$8.4 \pm 0.6 \times 10^{-17}$	$\gamma + \gamma$	98.8%	135
				$e^-e^+ + \gamma$	1.2%	

of magnitude shorter than that of its I-spin partners, the $\pi^\pm$. This decay is consistent with electromagnetic selection rules, and is not caused by the weak interaction.

In the quark model, the π^+ is composed of the $(u\bar{d})$ quark–antiquark pair. One could imagine the $q\bar{q}$ pair annihilating into a virtual W$^+$ that then decays into $\mu^+ + \nu_\mu$. An analogy would be positronium: $e^- + e^+ \to \gamma + \gamma$ with a lifetime 0.125 ns. Positronium is analyzed in a hybrid fashion in Chapter 9, in which the bound state is treated like an ordinary non-relativistic hydrogen atom with reduced mass $m_e/2$, and the annihilation is quantum electrodynamics. The rate is proportional to the square of the wave function at the origin. We do not have the ability to apply such machinery to the light quark bound states. So, our analysis of pion decay must be phenomenological, but it is nevertheless possible to make some progress. We shall see that the ratio $(\pi \to e\nu)/(\pi \to \mu\nu)$ can be unambiguously predicted.

For $\pi\mu$ decay, the T matrix element can be written

$$\left\langle \mu^+\nu_\mu \right| T \left|\pi^+\right\rangle = \frac{G_F \cos\theta_c}{\sqrt{2}} M^\lambda \bar{u}(p_2)\gamma_\lambda(1-\gamma^5)v(p_1);$$
$$q(\pi^+) = p_1(\mu^+) + p_2(\nu_\mu). \tag{6.60}$$

The lepton current is unambiguous. So, what is the four vector M^λ? The pion has spin zero, so there is no vector there. There is only the pion four momentum q^λ, so we postulate that:

$$M^\lambda = f_\pi(q^2) \times q^\lambda, \tag{6.61}$$

where $f_\pi(q^2)$ is an unknown invariant function with dimensions of mass. Since the pion is at rest, $q^\lambda = (m_\pi, 0, 0, 0)$; f_π is a constant, and $\lambda = 0$. Our T matrix is then:

$$\left\langle \mu^+ \nu_\mu \right| T \left| \pi^+ \right\rangle = \frac{G_F \cos\theta_c}{\sqrt{2}} f_\pi m_\pi \bar{u}(p_2)\gamma_0(1-\gamma^5)v(p_1), \tag{6.62}$$

and

$$\begin{aligned} \sum_{\text{spins}} |\left\langle f \right| T \left| i \right\rangle|^2 &= \frac{G_F^2 \cos^2\theta_c}{2} f_\pi^2 m_\pi^2 \operatorname{Tr}(\not{p}_2 \gamma_0 (1-\gamma^5)(\not{p}_1 - m_\mu)\gamma_0(1-\gamma^5)) \\ &= \frac{G_F^2 \cos^2\theta_c}{2} f_\pi^2 m_\pi^2 2\operatorname{Tr}(\not{p}_2 \gamma_0 (1-\gamma^5)\not{p}_1 \gamma^0). \end{aligned} \tag{6.63}$$

The term proportional to the muon mass does not contribute, nor does the γ^5 term, because of the two γ_0's, so we are left with only the vector part:

$$2\operatorname{Tr}(\not{p}_2 \gamma_0 \not{p}_1 \gamma_0) = 8(2p_2^0 p_1^0 - p_1 \cdot p_2). \tag{6.64}$$

This is of historical interest. Parity violation leads to the helicity rules, that forbid the decay $\pi \rightarrow l_1 + \bar{l}_2$, where both particles are massless leptons. Momentum balance requires $\vec{p}_1 = -\vec{p}_2$, and the helicities are opposite, so the parent particle cannot have spin zero. Equation (6.64) says that the γ^5 term drops out, so the vector coupling alone, without parity violation gives the same result. We have seen this phenomenon before, in the discussion of e^-e^+ annihilation, where the two spins add, so that in the center of mass the helicities are opposite. The vector interaction does not care whether it is a particle or an antiparticle that has negative helicity, and any conclusions drawn about the charged pion decay are the same for V and $V \pm A$. On the other hand, pure vector would not polarize the final state leptons.

To get a formula for the transition rate we need to evaluate the two-body phase space:

$$d\Gamma = \frac{G_F^2 \cos^2\theta_c f_\pi^2 m_\pi^2}{4m_\pi 4\pi^2} \left\{ 8(E_1 E_2 + \vec{p}_1 \cdot \vec{p}_2) \right\} \delta^{(4)}(q - p_1 - p_2) \frac{d^3 p_1}{2E_1} \frac{d^3 p_2}{2E_2}. \tag{6.65}$$

The neutrino is p_2, and the momenta are equal and opposite, so $E_2 = |\vec{p}_2| = |\vec{p}_1| = |\vec{p}|$, and $E_1E_2 + \vec{p}_1 \cdot \vec{p}_2 = |\vec{p}|(E_1 - |\vec{p}|)$. In a two-body decay all final state kinematics depend only on the masses.

$$E_1 = \frac{m_\pi^2 + m_\mu^2}{2m_\pi}; \quad E_2 = p = \frac{m_\pi^2 - m_\mu^2}{2m_\pi}. \tag{6.66}$$

The evaluation of the phase space proceeds normally. Integrate over the neutrino momentum:$\int \delta(\vec{p}_1 + \vec{p}_2)d^3p_2 = 1$. The energy term is

$$\int d^3p_1 \delta\left(m_\pi - E_1 - \sqrt{E_1^2 - m_\mu^2}\right) = \frac{4\pi p^2 E_1}{m_\pi}. \tag{6.67}$$

Using the energies in terms of the masses gives the final answer:

$$\Gamma = \frac{G_F^2 \cos^2\theta_c f_\pi^2}{8\pi} m_\pi m_\mu^2 \frac{(m_\pi^2 - m_\mu^2)^2}{m_\pi^4}. \tag{6.68}$$

All very nice, except that f_π is unknown. Whatever it is, it cancels in the ratio:

$$\frac{\Gamma(\pi \to e\nu)}{\Gamma(\pi \to \mu\nu)} = \frac{(m_\pi^2 - m_e^2)^2 m_e^2}{(m_\pi^2 - m_\mu^2)^2 m_\mu^2} = 1.28 \times 10^{-4}. \tag{6.69}$$

The ratio is about 4% larger than the experimental number quoted in Table 6.1. It is conventional wisdom that the suppression of the electron mode is a parity violation effect, but we have seen that this is not the case. The vector weak interaction does the suppression all by itself, but without $V - A$ there would be no muon polarization, and the muons from pion decay are 100% polarized, a property clearly demonstrated in Figure 6.4. Setting the unknown constant $f_\pi = 0.94m_\pi$ gives the correct pion lifetime, so f_π is very close to the pion mass.

The decay $\pi^+ \to \pi^0 + e^+ + \nu_e$ is analogous to ^{14}O decay, even including the $\sqrt{2}$ from I-spin, and the $\cos\theta_c$ for the u and d quarks. So, we begin with Equation (6.28), where now $Q = m_{\pi^+} - m_{\pi^0} = 4.59$ MeV, $a = m_e/Q = 0.111$, and the integral is 0.030. The approximation that the final state π^0 is infinitely heavy is only good to about 3%, but we make it anyway. Doing the arithmetic, and using the values of G_F and $\hbar$ quoted earlier, gives:

$$\begin{aligned}\Gamma(\pi^+ \to \pi^0 e^+ \nu_e) &= \frac{G_F^2 \cos^2\theta_c}{\pi^3} Q^5 \int_a^1 x(1-x)^2\sqrt{x^2-a^2}dx \\ &= 0.38 \text{ s}^{-1} \\ &= 1.02 \times 10^{-8} \times 0.38 \times 10^8. \end{aligned} \tag{6.70}$$

The last two numbers are the experimental values of the branching ratio and the reciprocal of the π^+ lifetime. The numbers agree within the precision of the calculation.

6.5. Weak Decays of Strange Particles

Almost all of the strange particles decay leptonically. $K^+ \to \mu + \nu_\mu$ is the analog of pion decay, and has similar problems, with an unknown parameter f_K. The baryons $\Lambda \to p\ e^-\bar{\nu}_e$ and $\Sigma^- \to n\ e^-\bar{\nu}_e$ are analogs of neutron decay, and can be analyzed in the same fashion. However, we will begin our discussion with the non-leptonic decays. These decay modes were the first to be discovered, and the puzzle of strong production versus weak decay was the motivation for the proposed new quantum number called ‘strangeness’. If the production process with a cross-section in mb $\pi^- + p \to \Lambda$, then why doesn’t $\Lambda \to \pi^- + p$ proceed with a lifetime characteristic of the strong interaction, $\sim 10^{-23}$ s, rather than the observed lifetime of about 10^{-10} s? A new quantum number came to the rescue. In Λ production the pion and the proton each have strangeness $S = 0$. The strong interaction conserves strangeness, so the final state is modified by adding a second strange particle, and assigning them opposite strangeness. Thus strange particle production is associated production; $S = +1$ and $S = -1$ particles are produced in pairs, and the weak interaction does not conserve strangeness.

$$\pi^- + p \to \Lambda + K^0; \quad \Lambda(S=-1) + K^0(S=+1);\ \text{decay in} \sim 10^{-10}\text{s};$$
$$\Lambda \to p\pi^-;\ \text{and}\ K^0 \to \pi^+\pi^-.$$

The lifetimes are in the range of 10^{-10} s. That this number is reasonable for the weak interaction can be argued on dimensional

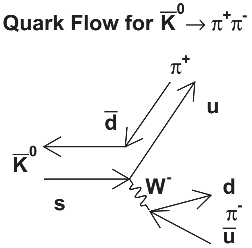


Figure 6.5: Quark flow diagram for $\bar{K}^0 \to \pi^+\pi^-$. The s quark emits a W^- and becomes a u quark. The $\bar{d}$ quark is a spectator. The W^- decays into $d\bar{u}$ that forms a π^-.

grounds. The exponents for G_F^2 and $\hbar$ are each 10^{-22}, and cancel. If $Q \sim 100\,\text{MeV}$, then $Q^5 = 10^{10}$, so $\tau \sim \hbar/(\mathrm{G_F^2}Q^5) \sim 10^{-10}\,\text{s}$. For $K \to \pi\pi$ one can draw a quark diagram like the one shown in Figure 6.5, in which an s quark becomes a u quark by emitting a W^- that decays into $d\bar{u}$. The $\bar{d}$ is a spectator. Final state $u\bar{d}$ and $d\bar{u}$ form the $\pi^+\pi^-$. Similar diagrams can be drawn for $\Lambda \to p\pi^-$, $\Sigma^+ \to p\pi^0$, and other non-leptonic decays. Every allowed decay can be represented by a connected quark diagram, in which the strange quark changes to a u quark. But unlike the Feynman diagrams of QED, these pictures do not furnish rules for calculation of decay rates, because of the recombination of the quarks to form mesons and baryons. The phase space factor in the K rest frame is

$$\int \frac{(2\pi)^4}{2m_K(2\pi)^6}\delta^4(p_k - p_1 - p_2)\frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2} = \frac{1}{16\pi m_K}\frac{p}{E}. \tag{6.71}$$

The velocity of either final state pion is $p/E = 205/249 = 0.82$. We may write the unknown matrix element of the weak Hamiltonian as follows:

$$\left\langle \pi^+\pi^- \right| H_W \left| K_S \right\rangle = \frac{\sin(\theta_c)G_F m_K^2 f_K}{\sqrt{2}}, \tag{6.72}$$

where f_K has the dimension of mass, like f_π in above formulas for pion decay. Note that $\sin\theta_c = 0.225$ appears in this formula. There is a brief discussion of the quark mixing matrix below, justifying the Cabibbo angle for the d and s quarks. Note also that the matrix element connects K_S^0 to $\pi\pi$. The mixing of the K^0 and $\bar{K}^0$ was discussed in Chapter 1, and the application of complex solutions of the Klein–Gordon equation to the $(K^0, \bar{K}^0) \leftrightarrow (K_S, K_L)$ was noted in Chapter 3. The conservation of hypercharge $Y = N+S$ replaces ordinary charge conservation to distinguish $K^0 = \Phi(x)$ from $\bar{K}^0 = \Phi^\dagger(x)$ in strong interactions, and the real fields $\Phi_1 = (\Phi + \Phi^\dagger)/\sqrt{2}$, and $\Phi_2 = (\Phi - \Phi^\dagger)/(i\sqrt{2})$ are K_S and K_L, respectively. The tiny mass difference $m_{K_L} - m_{K_S} = 3.5 \times 10^{-6}$eV is a symmetry breaking parameter. Interesting physical effects that follow from this decomposition, like strangeness oscillation, were discussed in Chapter 1. The decay that we are attempting to describe is $K_S \to \pi^+\pi^-$. Assembling the phase space and matrix element gives:

$$\Gamma(K_S \to \pi^+\pi^-) = \frac{\sin^2(\theta_c) G_F^2 m_K^4 f_K^2}{32\pi m_K} \times 0.82, \tag{6.73}$$

and the experimental numbers are

$$\frac{\Gamma}{\hbar} = 0.69 \times 1.12 \times 10^{10}\ \mathrm{s}^{-1}. \tag{6.74}$$

The first number in Equation (6.74) is the branching fraction for $K_S \to \pi^+\pi^-$, and the second number is the reciprocal of the K_S lifetime. Doing the arithmetic gives the unknown factor $f_K/m_k = 3.1$. The branching fraction for $K_S \to \pi^0\pi^0 = 0.31$, so the ratio $(\pi^+\pi^-)/(\pi^0\pi^0) \sim 2$. This number is predicted by the $|\Delta I| = 1/2$ rule discussed in Chapter 1. According to the rule, the two-pion final state, that could have $I = 2, 1, 0$ must have $I = 0$, because for two π^0's the Clebsch–Gordan coefficient for $I = 1$ vanishes (see Table 6.2), and I-spin should be the same for the two states. The coefficients for $I_1 = 1, I_2 = 1$, and $I_{\text{tot}} = 0$ are $\sqrt{1/3}, -\sqrt{1/3}$, and $\sqrt{1/3}$. The $\pi^+\pi^-$ gets two of them, and the $\pi^0\pi^0$ gets one, hence the ratio.

The final state of the non-leptonic decay $\Lambda \to p\pi^-$ can be analyzed using ordinary quantum mechanics. In the Λ rest frame spin $\uparrow$ along the z axis the proton angular momentum wave function is

$$\begin{aligned}\psi(1/2, +1/2) = S Y_0^0 \left|1/2, +1/2\right\rangle + P \Big(& Y_1^0 \left|1/2, +1/2\right\rangle / \sqrt{3} \\ & - \sqrt{2} Y_1^{+1} \left|1/2, -1/2\right\rangle / \sqrt{3} \Big),\end{aligned} \tag{6.75}$$

and for Λ spin $\downarrow$:

$$\psi(1/2,-1/2) = SY_0^0\,|1/2,-1/2\rangle - P\left(Y_1^0\,|1/2,-1/2\rangle/\sqrt{3}\right.$$
$$\left. -\ \sqrt{2}Y_1^{-1}\,|1/2,+1/2\rangle/\sqrt{3}\right). \qquad (6.76)$$

Here, S and P are unknown amplitudes for $l = 0$ and $l = 1$ in the final state, respectively; the kets on the right-hand side are proton spin functions, and the spherical harmonics describe the angles of the proton momentum vector. The Clebsch–Gordan coefficients for adding spin 1/2 to $l = 1$ are given in Table 1.2. The presence of both $l = 0$ and $l = 1$ final state amplitudes builds in parity violation. The S and P amplitudes are relatively real if time reversal invariance is valid. The final state interaction between the proton and π^- adds a phase difference $\delta_s - \delta_p = 6.5°$, and Λ decay measurements are consistent with this being the only phase. The magnitudes $\psi^\dagger\psi$ are

$$|\psi_+|^2 = \frac{(S^2+P^2)}{4\pi}(1+\alpha\cos\theta),\ \text{where}, \alpha = \frac{2SP}{S^2+P^2}, \qquad (6.77)$$

and

$$|\psi_-|^2 = \frac{(S^2+P^2)}{4\pi}(1-\alpha\cos\theta). \qquad (6.78)$$

The angle θ is the polar angle of the proton relative to the $+z$ axis in the Λ rest frame. If the parent Λ polarization $\langle\sigma_z\rangle = P_\Lambda < 1$, then $\alpha \to \alpha\ P_\Lambda$. For unpolarized hyperons there is no asymmetry. However, the daughter protons are longitudinally polarized. A fairly simple calculation is $\langle\sigma_z\rangle$ for the final state proton.

$$\sigma_z\psi_+ = SY_0^0\,|1/2,+1/2\rangle + P\left(Y_1^0\,|1/2,+1/2\rangle/\sqrt{3}\right.$$
$$\left. +\ \sqrt{2}Y_1^{+1}\,|1/2,-1/2\rangle/\sqrt{3}\right). \qquad (6.79)$$

Multiplying from the left by $\psi_+^\dagger$ and dividing by the normalization Equation (6.77) gives the expectation value:

$$\langle\sigma_z\rangle = \frac{S^2 - P^2 + 2SP\cos\theta + 2P^2\cos^2\theta}{(S^2+P^2)(1+\alpha\cos\theta)}. \qquad (6.80)$$

Defining the parameter $\gamma = (S^2 - P^2)/(S^2 + P^2)$, this formula can be written:

$$\langle\sigma_z\rangle = \frac{\gamma + (1-\gamma)\cos^2\theta + \alpha\cos\theta}{1+\alpha\cos\theta}; \ \Lambda \text{ spin } \uparrow. \tag{6.81}$$

For Λ spin $\downarrow$ we have

$$\langle\sigma_z\rangle = \frac{-\gamma - (1-\gamma)\cos^2\theta + \alpha\cos\theta}{1-\alpha\cos\theta}; \ \Lambda \text{ spin } \downarrow. \tag{6.82}$$

The common term $\alpha\cos\theta$ that persists is a longitudinal polarization of the proton from unpolarized Λ's. The coefficient $\gamma = (S^2 - P^2)/(S^2 + P^2)$ measures the amount of parent Λ polarization that is transferred to the daughter proton. Note that $\gamma = \pm 1$ for $P = 0$ or $S = 0$ respectively, so that, unlike α, it exists if parity is conserved. If $\gamma = 1$, pure S wave, then $\langle\sigma_z\rangle = 1$. The Particle Data Group [5] gives the following formula for the proton polarization:

$$\vec{P}_B = \frac{(\alpha + \vec{P}_Y\cdot\hat{n})\hat{n} + \beta(\vec{P}_Y\times\hat{n}) + \gamma\hat{n}\times(\vec{P}_Y\times\hat{n})}{1+\alpha\vec{P}_Y\cdot\hat{n}}; \tag{6.83}$$

where the proton momentum direction in the Λ rest frame is $\hat{n}$, and $\vec{P}_Y$ is the Λ spin polarization vector. The term $\beta = 2\mathrm{Im}(S^*P)/(|S|^2 + |P|^2)$ violates time reversal invariance, and does not contribute to our formulas for $\langle\sigma_z\rangle$. Setting $\vec{P}_Y = \hat{z}$, and calculating $\vec{P}_B\cdot\hat{z}$ gives Equation (6.81). The same analysis applies to the non-leptonic decay of any of the spin 1/2 baryons. The $|\Delta I| = 1/2$ rule makes several predictions for Λ and Σ decay, that are covered in the Problems to this Chapter. The Ω^- has spin 3/2, so the interference is between $l = 1$ and $l = 2$, with different Clebsch–Gordan coefficients and a different angular dependence. The asymmetry parameter α is different for the various Ω^- decay modes, but is small in each case. See Table 6.3.

Longitudinal polarization of daughter baryons gives a technique for measuring the asymmetry parameter α, because as noted above the proton or pion asymmetry is the product αP_Λ, and the Λ polarization that arises from strong interaction in a reaction like $\pi^- + p \to \Lambda + K^0$, while normal to the production plane, is usually unknown in either sign or magnitude. Longitudinal proton polarization is not so easy to measure either, but these formulas are all in the

Table 6.3: Properties of the baryons.

Symbol	Mass (MeV)	Spin/ parity	Lifetime (s)	Decay	$\Gamma_i/\Gamma_{\text{tot}}(\%)$	Q (MeV)	α parameter
Λ	1115.7	$1/2^+$	2.63×10^{-10}	$p\pi^-$	64%	37	$+0.750 \pm 0.009$
				$n\pi^0$	36%	41	$+0.75 \pm .05$
Σ^+	1189.4	$1/2^+$	0.80×10^{-10}	$p\pi^0$	51.6%	116	-0.98 ± 0.02
				$n\pi^+$	48.3%	110	0.068 ± 0.013
Σ^-	1197.4	$1/2^+$	1.48×10^{-10}	$n\pi^-$	99.8%	118	-0.068 ± 0.008
Ξ^0	1315	$1/2^+$	2.9×10^{-10}	$\Lambda\pi^0$	∼100%	65	-0.35 ± 0.01
Ξ^-	1321	$1/2^+$	1.64×10^{-10}	$\Lambda\pi^-$	∼100%	66	-0.392 ± 0.008
Ω^-	1672	$3/2^+$	0.82×10^{-10}	ΛK^-	68%	63	-0.026 ± 0.026
				$\Xi^0\pi^-$	24%	217	0.09 ± 0.14
				$\Xi^-\pi^0$	8.6%	216	0.05 ± 0.21

Λ rest frame, and Lorentz transformation into the laboratory comes to the rescue. While for a massless Fermion the helicity is Lorentz invariant (see Problem 5.12), this is not true for protons from Λ decay, where the proton velocity is only $v \sim 0.1$. The proton spin stays fixed to within a few degrees as in the Λ rest frame, while the proton momentum vector is swept forward in the direction of the parent Λ. In this way, a transverse proton polarization is created, that could be measured by a right–left asymmetry for proton scattering on carbon. The various angles and transformations are described in the experimental paper of Cronin and Overseth [6]. This technique has been superseded by e^-e^+ colliding ring experiments that studied the reaction $e^-e^+ \rightarrow \Lambda\bar{\Lambda}$ both on the J/ψ resonance and off of it [7].

Table 6.3 summarizes the parameters of the strange baryons that decay by weak interaction into a baryon and a meson. Q is the kinetic energy available to the final state. The lifetimes are all about the same. Note that α_- for $\Lambda \rightarrow p\pi^-$ has increased from previous listings by 17% from 0.64 to 0.75, from the BESIII reference above [7]. This has changed several asymmetry numbers in the table. For example if the Ξ^0 sample has zero polarization, the daughter Λ is longitudinally polarized by α_{Ξ_0}. This effect is in turn measured by the decay proton asymmetry from $\Lambda \rightarrow p\pi^-$, so that the measured parameter is $\alpha_{\Xi_0}\alpha_-$. For the decay $\Lambda \to n\pi^0$ the measured quantity was the ratio α_0/α_-. Numbers in this table have been updated to account for this change. There have been many strong interaction production experiments that directly measure the Λ polarization by observing the decay asymmetry. Parity violation in the strange baryons means that they are their own polarization analyzers. The increase in α_- means that all of these polarizations are 17% smaller than they were thought to be before.

The scattering amplitude $f(\theta)$ given in Equation (1.5) is for spinless particles, and has to be modified in order to describe spin effects in scattering reactions. We limit the discussion to one spin 1/2 particle and a spin zero target. The quantization axis $\hat{z}$ is in the incident beam direction, as was the case for Equation (1.5). The spin vectors $|1/2, +1/2\rangle$ and $|1/2, -1/2\rangle$ are oriented along the beam. The scattering plane is the (x, z) plane. Quantum mechanics gives us the opportunity to create a polarization normal to the scattering plane by judicious choice of the mixture of spin ±1/2 kets along the beam. Parity conservation in strong interactions precludes a dependence of the cross-section on whether the incident spin 1/2 particle has helicity +1 or −1. Transverse spin can be flipped by a spin orbit force of the type $U(r)\vec{L}\cdot\vec{S}$ that occurs in atomic physics. In Chapter 4, we discussed the use of Compton scattering to measure photon polarization; that was a parity conserving electromagnetic interaction, but there were two spins, the electron and the photon, that could be manipulated. Here, there is only one spin 1/2 particle, and it is not possible for $\vec{L} = \vec{r}\times\vec{p}$ to have a component along the beam.

We will follow the analysis of Merzbacher [8], for a description of spin dependent scattering, although the subject is treated in many standard texts on quantum mechanics. First we introduce a density matrix to treat spin 1/2 polarization. An arbitrary spin wave function is written

$$\chi = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} |c_1|e^{i\gamma_1} \\ |c_2|e^{i\gamma_2} \end{pmatrix}, \ \chi^\dagger = \begin{pmatrix} |c_1|e^{-i\gamma_1} & |c_2|e^{-i\gamma_2} \end{pmatrix},$$

$$\chi^\dagger\chi = |c_1|^2 + |c_2|^2 = 1, \tag{6.84}$$

where it is only the phase difference $\gamma_1 - \gamma_2$ that has physical significance. The density matrix $\rho = \chi\chi^\dagger$:

$$\rho = \chi\chi^\dagger = \begin{pmatrix} |c_1|^2 & c_1 c_2^* \\ c_1^* c_2 & |c_2|^2 \end{pmatrix}. \tag{6.85}$$

For an observable in spin language A (like one of the spin components, say σ_x), the expectation value for the state χ is

$$\begin{aligned} \langle A \rangle &= \chi^\dagger A \chi = \begin{pmatrix} |c_1|e^{-i\gamma_1} & |c_2|e^{-i\gamma_2} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} |c_1|e^{i\gamma_1} \\ |c_2|e^{i\gamma_2} \end{pmatrix} \\ &= \mathrm{Tr}(\rho A) = \mathrm{Tr}(A\rho). \end{aligned} \tag{6.86}$$

A density matrix is usually applied to an ensemble of randomly oriented spins. The state defined by χ is not random. In fact, $\langle \vec{\sigma} \cdot \hat{n} \rangle = 1$, where the unit vector $\hat{n} = \cos\theta\hat{z} + \sin\theta\cos\phi\hat{x} + \sin\theta\sin\phi\hat{y}$, and the components of χ are $|c_1|e^{i\gamma_1} = \cos(\theta/2)e^{-i\phi/2}$ and $|c_2|e^{i\gamma_2} = \sin(\theta/2)e^{i\phi/2}$. A completely random unpolarized spin sample would be described by the diagonal density matrix $\rho = 1/2$, for then $\mathrm{Tr}\rho\sigma_i = 0$ for all three components of σ_i. Our density matrix can be written:

$$\begin{aligned} &\rho = \frac{1}{2}(1 + \vec{\sigma} \cdot \hat{n}), \text{ where } \vec{\sigma} \cdot \hat{n} = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix} \\ &\text{and } \mathrm{Tr}(\rho\vec{\sigma} \cdot \hat{n}) = 1. \end{aligned} \tag{6.87}$$

We return to the spin dependent scattering problem. The scattering amplitude now has two components:

$$\begin{pmatrix} \psi_1(\theta,\phi) \\ \psi_2(\theta,\phi) \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} e^{ikz} + \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \frac{e^{ikr}}{r}. \tag{6.88}$$

Here, c_1 and c_2 are complex numbers describing the incident polarization, and S is the scattering matrix, with components that depend on the angles θ and ϕ. The component of total angular momentum J_z along the incident beam direction is $+\hbar/2$ for ψ_1 and $-\hbar/2$ for ψ_2, and these components must be preserved in the scattering. Thus, we write

$$\hbar\left(\frac{-i\partial}{\partial\phi}+\frac{\sigma_z}{2}\right)\psi_1(\theta,\phi)=\hbar\left(\frac{c_1}{2}e^{ikz}+\left(\frac{S_{11}c_1}{2}-\frac{S_{12}c_2}{2}\right.\right.$$
$$\left.\left.-\frac{i\partial}{\partial\phi}(S_{11}c_1+S_{12}c_2)\right)\frac{e^{ikr}}{r}\right). \quad (6.89)$$

The right-hand side equals $\hbar\psi_1/2$ if $\partial S_{11}/\partial\phi = 0$, and $\partial S_{12}/\partial\phi = iS_{12}$. So, we may write

$$S_{11}=g(\theta), \quad \text{and} \quad S_{12}=h(\theta)e^{i\phi}, \quad (6.90)$$

where $g(\theta)$ and $h(\theta)$ are unknown functions of the polar scattering angle that could in principle be solved from the Schrödinger equation, given the potentials. Note that the phase of S_{12} is just what is needed to keep the total magnetic quantum number $+1/2$. This analysis can be repeated for the lower component:

$$\hbar\left(\frac{-i\partial}{\partial\phi}+\frac{\sigma_z}{2}\right)\psi_2(\theta,\phi)=\hbar\left(-\frac{c_2}{2}e^{ikz}+\left(\frac{S_{21}c_1}{2}-\frac{S_{22}c_2}{2}\right.\right.$$
$$\left.\left.-\frac{i\partial}{\partial\phi}(S_{21}c_1+S_{22}c_2)\right)\frac{e^{ikr}}{r}\right). \quad (6.91)$$

Setting this equal to $-\hbar\psi_2/2$ gives:

$$S_{22}=g(\theta), \quad \text{and} \quad S_{21}=-h(\theta)e^{-i\phi}. \quad (6.92)$$

Note that the functions $g(\theta)$ for S_{11} and S_{22} are the same. Parity conservation precludes any dependence on the helicity of the initial state. For the off diagonal elements the parity operation sends $\phi\rightarrow\pi+\phi$, and $e^{i\pi}=-1$, hence the minus sign on S_{21} (it has to be on one off diagonal element). Assume the scattering takes place in the (x,z) plane, so that $\hat{y}$ is normal to the scattering plane, and the incident beam is 100% polarized along the $\hat{y}$ direction. Then:

$$\rho=\frac{1+\sigma_y}{2}=\frac{1}{2}\begin{pmatrix}1 & -i\\ i & 1\end{pmatrix}, \quad \text{and}$$
$$S^\dagger S=\begin{pmatrix}|g|^2+|h|^2 & 2i\mathrm{Im}(g^*h)e^{i\phi}\\ -2i\mathrm{Im}(g^*h)e^{-i\phi} & |g|^2+|h|^2\end{pmatrix}. \quad (6.93)$$

The polarized cross-section is

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \mathrm{Tr}(\rho S^\dagger S) = \frac{1}{2}(2(|g|^2 + |h|^2) - 2\mathrm{Im}(g^* h)(e^{i\phi} + e^{-i\phi})) \\ &= (|g|^2 + |h|^2)(1 + \alpha \cos\phi),\end{aligned} \tag{6.94}$$

where the asymmetry coefficient (that would be multiplied by $P < 1$ for a partially polarized beam) is

$$\alpha = -\frac{2\mathrm{Im}(g^* h)}{|g|^2 + |h|^2}. \tag{6.95}$$

The azimuthal angle ϕ is measured relative to the x axis, and this is called a 'left–right' asymmetry. Now suppose the same scattering is done with an unpolarized incident beam. It suffices to assume only $c_1 = 1$, and $c_2 = 0$, a right-handed helicity. The scattering matrix S is the same. The scattered spin function is

$$\begin{pmatrix} g(\theta) & h(\theta)e^{i\phi} \\ -h(\theta)e^{-i\phi} & g(\theta) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} g(\theta) \\ -h(\theta)e^{-i\phi} \end{pmatrix} \text{ and } \langle \sigma_y \rangle = \frac{-2\mathrm{Im}(g^* h)}{|g|^2 + |h|^2}, \tag{6.96}$$

for $\phi = 0$, that is scattering to the left. The polarization changes sign for scattering to the right. This formula is the same as the the asymmetry in scattering of a 100% polarized beam, and gives a prescription for determining the magnitude of an unknown polarization by performing a double scattering experiment. The first scattering is the polarizer, and the second is the analyzer, and the final left–right asymmetry is the magnitude squared of the polarization. You have to set the same polar angle for each scatter, and the same scattering plane. So, this technique gives you the magnitude by taking the square root of the asymmetry, but you do not know the sign. Polarization from scattering is not a weak interaction, but because of parity violation the weak interaction leads to many polarized spin 1/2 particles that have historically been analyzed by exploiting the technique derived here. ^{12}C is a spin zero nucleus, so this formalism can treat the scattering of protons by carbon nuclei.

We now turn to the beta decay of the Λ as a final topic in the weak decays of strange particles. The Λ lifetime is $\tau_\Lambda = 2.63 \times 10^{-10}$ s, and the branching fraction for $\Lambda \to p e^- \bar{\nu}_e = 8.3 \pm 0.14 \times 10^{-4}$, giving

$\Gamma(\Lambda \to pe^-\bar{\nu}_e) = 3.15 \times 10^6\,\mathrm{s}^{-1}$. We attempt to describe this decay with a variation of the neutron formula. We adopt the T matrix element from Equation (6.3), with the addition of an SU(3) factor that couples $\langle\Lambda|\, V_-\, |p\rangle = \sqrt{3/2}$. (This is the same matrix element as $\langle\eta^0|\, V_-\, |K^+\rangle$ in Table 1.4.)

$$\begin{aligned}\langle pe^-\bar{\nu}_e|\, T\, |\Lambda\rangle = {} & \frac{G_F \sin(\theta_c)}{\sqrt{2}}\sqrt{3/2}(\bar{u}_p\gamma^\mu(1-\gamma^5)u_\Lambda) \\ & \times (\bar{u}_e\gamma_\mu(1-\gamma^5)v_{\nu_e}). \end{aligned} \tag{6.97}$$

The $\cos(\theta_c)$ term has been replaced with $\sin(\theta_c)$, because of the strange quark transition $s \to u$, and the V spin matrix element is like the $\sqrt{2}$ from I-spin that appeared in the ^{14}O formula. Equation (6.25) is similarly modified, in addition to setting $|v_e| = 1$, because the energy available $Q = 177\,\mathrm{MeV}$, and the electron is fully relativistic. It may be a bit of a stretch to assume that the final state proton is at rest, but the maximum kinetic energy is only 14 MeV. So with the caveat that this may not be precisely correct, we use the two body phase space from neutron decay. The parameter λ introduced on the axial vector term of neutron decay is known for Λ beta decay: $\lambda = 0.72$.

$$\begin{aligned}\sum_{\text{spins}} |\langle pe^-\bar{\nu}_e|\, T\, |\Lambda\rangle|^2 = {} & \frac{3G_F^2\sin^2(\theta_c)}{4} 64 m_p m_\Lambda E_e E_\nu (1+3\lambda^2) \\ & \times \left(1 + \frac{1-\lambda^2}{1+3\lambda^2}\cos(\theta_{e\nu})\right). \end{aligned} \tag{6.98}$$

Because of the large Q value we can set $a = 0$ in Equation (6.29), giving the phase space integral:

$$\int_0^1 x^2(1-x)^2 dx = \frac{1}{30}. \tag{6.99}$$

So, we have the decay rate:

$$\Gamma(\Lambda \to pe^-\bar{\nu}_e) = \frac{3G_F^2 \sin^2\theta_c Q^5 (1+3\lambda^2)}{120\pi^3\hbar} = 3.7 \times 10^6\,\mathrm{s}^{-1}. \tag{6.100}$$

This number is 15% larger than the experimental value, but within about one standard deviation of the measured branching ratio. Without the effect of the Cabbibo angle it would be too large by a factor of 20. It was this discrepancy that motivated Cabibbo's idea of quark mixing, before there were quarks! There was also an unexplained difference between neutron decay and muon decay that was fixed by $\cos\theta_c$ [9].

6.6. Hadrons with Charm and Bottom Quarks

While strangeness can be treated without the quark model, it is virtually impossible to understand the heavier D and B mesons without quarks. The electroweak theory is based on quarks and leptons coupling to weak intermediate bosons. The six leptons are arranged in left-handed doublets of 'weak I-spin', with the neutrinos having $I_z^W = +1/2$:

$$\begin{pmatrix} I_z^W = +1/2, q = 0 \\ I_z^W = -1/2, q = -1 \end{pmatrix}; \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}; \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}; \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}. \tag{6.101}$$

There are separately conserved lepton numbers for each doublet, assigned $l = +1$ for leptons, and $l = -1$ for antileptons, for which all quantum numbers are reversed. Weak transitions occur within one doublet; there is no mixing. The charge rule is $q = I_z^W - l/2$. In the lepton doublets the spinors are represented by the particles, and $e_L = (1 - \gamma^5)u(p_e)/2$. The right-handed components of the charged leptons $e_R = (1 + \gamma^5)u(p_e)/2$ are assigned weak I-spin zero. These components are absent from the weak interaction, but are present for electrodynamics. The neutrinos are assumed to be massless. Experimental demonstration of neutrino mixing means that this assumption is incorrect, although the masses are less than one eV. In fact, a recent direct measurement of $m_{\bar{\nu}e} \leq 1.1$ eV (90% confidence) has been reported by the Katrin Collaboration from the end point electron energy from the beta decay of tritium [10].

There are six quarks with fractional charge. Their masses were listed in Table 1.5. Some other relevant quantum numbers are given

Table 6.4: The quarks.

	d	u	s	c	b	t
Q	$-1/3$	$+2/3$	$-1/3$	$+2/3$	$-1/3$	$+2/3$
I-spin	$1/2$	$1/2$	0	0	0	0
I_z	$-1/2$	$+1/2$	0	0	0	0
S strange	0	0	-1	0	0	0
C charm	0	0	0	$+1$	0	0
B bottom	0	0	0	0	-1	0
T top	0	0	0	0	0	$+1$

in Table 6.4. Note that each quark past u and d have flavor quantum numbers. The first one was strangeness, assigned before the quark model, and with the convention originally established, the s quark has $S = -1$. This pattern is replicated for the heavier quarks, where the sign of the flavor quantum number is the same as the charge. Strong and electromagnetic interactions are assumed to conserve the flavor quantum number. The charge rule is $Q = I_z + (N + S + C + B + T)/2$, N being the baryon number. In the electroweak theory, these quarks are arranged in weak I-spin doublets. Each quark is assigned baryon number $N = 1/3$, because the octet baryons contain three quarks.

$$\begin{pmatrix} I_z^W = +1/2, q = +2/3 \\ I_z^W = -1/2, q = -1/3 \end{pmatrix}; \begin{pmatrix} u_L \\ d_L \end{pmatrix}; \begin{pmatrix} c_L \\ s_L \end{pmatrix}; \begin{pmatrix} t_L \\ b_L \end{pmatrix}. \tag{6.102}$$

The notation is the same as for the leptons. The quark spinors are all left handed in the relativistic limit. The right-handed quarks, like the right-handed charged leptons, are assigned weak I-spin zero. The projection operator $(1-\gamma^5)/2$ has little effect on most top quark spinors, since the top quark mass $m_t = 174\,\text{GeV}$. The up and down quarks have ordinary I-spin $1/2$ as well, while the other four quarks have I-spin 0. The intermediate bosons $W^\pm$ have unit charge, so a weak process like $c \to s + W^+$ conserves charge. There are six antiquarks as well, with opposite quantum numbers, giving $\bar{c} \to \bar{s} + W^-$. Transitions within a weak I-spin doublet are allowed, like $u \to d + W^+$ or $c \to s + W^+$, while transitions outside of the doublets are suppressed. The farther outside you go, the greater the suppression. Thus, $t \to b$ is allowed, while $t \to s$ is suppressed, and $t \to d$ is suppressed even more! This phenomenon is the basis for the Cabibbo angle $\theta_c = 13^\circ$. The allowed transition $u \to d + W^+$ is multiplied by $\cos\theta_c = 0.974$, while $u \to s + W^+$ gets $\sin\theta_c = 0.225$, suppressing the rate by a factor of 1/20. This is a mixing of the d and s quarks for the weak interaction. No such mixing occurs for the leptons. The charge +2/3 quarks in the heavy quark doublets (c and t) are much heavier than their charge −1/3 companions. One consequence of this is that the B mesons that contain the $\bar{b}$ quark must decay into $\bar{c}$ quarks outside their doublet, and therefore have relatively long lifetimes for their heavy mass ~4.5 GeV. This long lifetime has led to very fruitful experiments involving B mesons. With six quarks the mixing is expanded into a 3×3 unitary matrix called the CKM matrix, after Cabibbo, Kobayashi, and Maskawa. We will introduce the matrix eventually. It is complicated, but it uses trigonometric functions to achieve allowed, forbidden, and even more forbidden transitions, and it includes a complex phase to account for the CP violation mentioned in Chapter 1. In 2×2 form, the matrix is not large enough to incorporate a complex phase. We would write the quarks d' and s' that couple to the weak interacting u quark as follows:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \tag{6.103}$$

and $\bar{u}d' = \bar{u}d\cos\theta_c + \bar{u}s\sin\theta_c$.

The electroweak theory begins with a weak I-spin triplet of intermediate bosons, called $(W_\mu^+, W_\mu^0, W_\mu^-)$, and a weak I-spin scalar field called B_μ. W^0 and B fields then mix via the Weinberg angle θ_W (which was introduced by Sheldon Glashow) to form the neutral fields Z_μ and A_μ (the photon). All of the vector fields are massless to begin with, and three of them (W^+, W^-, and Z) acquire mass via dynamical symmetry breaking, a field theory manipulation. The neutral weak intermediate vector boson Z_μ has a mass slightly greater than the W, $m_Z = 91.1$ GeV. The Z couples to e^-e^+, $\mu^-\mu^+$, $\tau^-\tau^+$ pairs, and $\nu\bar{\nu}$ pairs of all three lepton flavors. It also couples to $q\bar{q}$ quark pairs of the same flavor, like $u\bar{u}$ or $b\bar{b}$, but not to $s\bar{d}$, a 'flavor changing neutral current'. It is an experimental fact that flavor changing neutral currents are strongly suppressed in nature. The first order electroweak six quark, six lepton theory forbids flavor changing neutral currents.

The D mesons all contain the charm quark c (see Table 6.5). There is a family of charge conjugate mesons, D^-, $\bar{D}^0$, that contain the $\bar{c}$ quark, and have charge conjugate decay modes. Because each D meson has about 50 different decay modes, with a branching fractions of a few percent or less, both inclusive and some of the exclusive decays are listed. They are characterized by very large Q values compared to strange and non-strange particle decays, and very short lifetimes $\sim 10^{-12}\,\mathrm{s}$. Short lifetimes mean short travel distances before decay. Without any help from time dilation, the average flight distance is $c\tau = 0.3\,\mathrm{mm}$. Because the D mass is $\sim 2\,\mathrm{GeV}$, the Lorentz factor $\gamma = E/m$ does not appreciably lengthen the decay path. The LHC at CERN has multi-TeV proton beams, and the LHCb experiment is built to observe heavy mesons in the forward direction, where $\gamma \sim 100$, but even so the decay length is only 3 cm. While there have been K meson beams ($\tau \sim 10^{-8}\,\mathrm{s}$) and hyperon beams ($\tau \sim 10^{-10}\,\mathrm{s}$) for decades, beams of D's or B's for detailed study of these heavy mesons do not exist. A significant amount of information regarding D meson decay has come from the daughters of B meson decay.

In the spectator model of D or B meson decay the light quark plays no role, and the decay $D^0 \rightarrow K^-\pi^+$ is simply $c \rightarrow sW^+$, with the final state quarks rearranging themselves at no cost. In this model, the lifetimes of the D^0 and D^+ mesons would be equal, but they differ by a factor of two. Table 6.6 shows that this idea works better for the B's, where the heavy quark mass is a factor of three or so larger.

The charm and strange quarks are in the same weak I-spin doublet, so the decay is allowed. Scaling the muon lifetime as $(m_\mu/m_D)^5$ gives an order of magnitude guess for $\tau_D \sim 0.3 \times 10^{-12}\,\mathrm{s}$. Note that the assignments of particles and antiparticles as characterized by the quark content is reversed in the D's compared to the K's. This is a consequence of the sign of the quark flavor following the sign of the charge, as mentioned above. The D^0 decays into a $\bar{K}^0$, and $\bar{K}^0$'s and K^-'s are characteristic of charm particle decays. The D_{s} mesons

Table 6.5: Charmed quark mesons.

P scalars (0^-) symbol	Mass (MeV)	Charge	Lifetime (s)	Decay	$\Gamma_i/\Gamma_{\text{tot}}$ (%)	Q (MeV)	Quark content
c quark	1275	+2/3	10^{-12}	$c \to s$	100	1180	c
D^+	1869.6	+1	1.04×10^{-12}	$e^+ + x$	16	—	$c\bar{d}$
				$K^- + x$	26	—	
				$\bar{K}^0 + x; K^0 + x$	61	—	
				$K^-\pi^+\pi^+$	9	1092	
				$\bar{K}^0 + e^+\nu_e$	8.7	1373	
D^0	1864.8	0	0.415×10^{-12}	e^+x or μ^+x	13	—	$c\bar{u}$
				$K^- + x$	55	—	
				$\bar{K}^0x + K^0x$	47	—	
				$K^-e^+\nu_e$	3.5	1371	
				$K^-\pi^+\pi^0$	14.4	1090	
D_s^+	1968.3	+1	0.50×10^{-12}	$\phi + x$	16 ± 1	—	$c\bar{s}$
				$K^+K^-\pi^+$	5.4 ± 0.2	838	
				$\phi l^+\nu_l$	2.0 ± 0.5	948	
Vectors (1^-)							
D^{*+}	2010	+1	$\Gamma < 0.1$ MeV	$D^0\pi^+$	68	5	$c\bar{d}$
				$D^+\pi^0$	31	6	
D^{*0}	2007	0	$\Gamma < 2.1$ MeV	$D^0\pi^0$	62	8	$c\bar{u}$
				$D^0\gamma$	38	142	
C Baryon							
Λ_c^+	2286	+1	0.21×10^{-12}	$\Lambda\pi^+\pi^0$	3.6 ± 1.3	890	udc

Table 6.6: Bottom quark mesons.

P scalars 0^- symbol	Mass (MeV)	Charge	Lifetime (s)	Decay	$\Gamma_i/\Gamma_{\text{tot}}$ (%)	Q (MeV)	Quark content
b quark	4180	$-1/3$	10^{-12}	$b \to c$			
B^+	5279.3 ± 0.1	$+1$	1.63×10^{-12}	$\bar{D}^0 + x$	79 ± 4	—	$u\bar{b}$
				$l^+\nu_l x$	11.0 ± 0.3	—	
				$\bar{D}^0 l^+ \nu_l$	2.35 ± 0.09	3415	
				$J/\psi K^+$	0.1	1684	
B^0	5279.6 ± 0.1	0	1.52×10^{-12}	$l^+\nu_l x$	10.3 ± 0.3	—	$d\bar{b}$
				$\bar{D}^0 x$	47.4 ± 2.8	—	
				$J/\psi K^0$	0.08	3415	
B_s^0	5367	0	1.5×10^{-12}	$D_s^- x$	93 ± 23	—	$s\bar{b}$
				$D_s^+ D_s^-$	0.4	1431	
B_c^+	6275	$+1$	0.5×10^{-12}	$J/\psi \pi^+$	seen	3035	$c\bar{b}$
Vector 1^-							
B^*	5325	$0/+1$	—	$B\gamma$	100	46	$u/d\bar{b}$
B baryon							
Λ_b	5620	0	1.47×10^{-12}	$\Lambda_c^+ l^- \bar{\nu}_l$	6.2 ± 1.4	3334	udb
				$\Lambda_c^+ D_s^-$	1.1 ± 0.1	1366	

are interesting because they contain an $\bar{s}$ quark in place of a $\bar{d}$. This means that the $D_\mathrm{s} \to \bar{K}K$, and the ϕ vector meson, that contains the $\bar{s}$s quark pair, appears prominently in the final state. The D^*'s are vector mesons that decay by strong interaction into $D + \pi$, so that a width in MeV is quoted rather than a lifetime. Only the charmed baryon Λ_c^+ is listed, although there are seven known: $\Lambda_c^+ = udc$, $\Sigma_c^{++} = uuc$; $\Sigma_c^+ = udc$; $\Sigma_c^0 = ddc$; $\Xi_c^+ = usc$; $\Xi_c^0 = dsc$; and $\Omega_c^0 = ssc$, all with spin/parity $(1/2)^+$ and $C = +1$.

The value of the quark model for classifying these heavy hadronic states is obvious. Note that the charged and neutral B mesons have nearly the same mass and lifetime. The energy available to generic semileptonic decays is calculated assuming that the charged lepton mass is zero. This is certainly safe for electrons, and perhaps for muons, but not for τ's; so for a $\tau\nu_\tau$ final state, subtract 1777 MeV. Hadrons containing the b quark have about the same weak decay lifetimes as those containing the c quark, about 10^{-12} s, although B masses are almost three times as large as D masses. The suppression of the decay rate comes from the CKM matrix, and the fact that $b \to c$ is outside the weak I-spin doublet. There are literally hundreds of decay modes for the B mesons, and the B baryons. The mass is so large that many particles can easily be accommodated in the final state. There is also an experimental problem related to sorting out the B decay data. Most of our knowledge of B mesons comes from e^-e^+ single ring colliders, where the $b\bar{b}$ resonant state called the upsilon (Υ) 4S, with $J^{PC} = 1^{--}$ and a mass of 10,580 MeV, decays into a $B\bar{B}$ pair at rest in the laboratory (22 MeV total kinetic energy available). Half of the pairs are $B^0\bar{B}^0$, and half are B^+B^-. Since everything is taking place at the annihilation vertex, it is not easy to sort out what is going on. The slow moving $B\bar{B}$ system helps in acceptance efficiency of all final state particles, and conservation of energy and charge are useful constraints. Almost all of the branching fractions are less than 1%, and the high multiplicity leads to many low energy particles that must be accounted for. Also, the Υ 4S rides on a continuum background of light quarks from $e^-e^+ \to q\bar{q}$. For all of these reasons these studies are experimentally challenging. These problems are reflected in the tables published by the Particle Data

Group, and referenced above. CLEO was one of the groups working at e^-e^+ storage rings that studied B decays and documented their experimental results. A typical complete paper is Asner *et al.* [11]. This paper describes exclusive measurements of semileptonic decays of the rare type $\mathrm{B}^0 \to \pi^- l^+ \nu_l$. This decay is $\bar{b} \to \bar{u}$, and hence a twice forbidden transition.

Quark flow diagrams are fun to draw, and although there are no Feynman rules for calculating cross-sections and decay rates from these pictures, one can learn something about favored and disfavored processes from quark flow. Hadrons that decay by strong interaction have smoothly flowing diagrams, like Figure 6.6 that shows the decay $\Delta^{++} \to p\pi^+$. Note that antiquarks flow backwards. Figure 6.7 shows the Υ 4S $b\bar{b}$ state, that is above threshold for meson production, smoothly going into $B^0\bar{B}^0$ pair. Figure 6.8 shows the decay of the ψ', that is below threshold for D meson production: $\psi' \to \mathrm{J}/\psi\pi^+\pi^-$. This diagram is disconnected, and the decay rate is lower (the state is narrower) than it would be if the ψ' could energetically decay

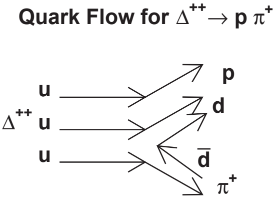


Figure 6.6: Quark flow diagram for the strong decay $\Delta^{++} \to p\pi^+$. No quarks change flavor. One u quark is a spectator, the other two combine with a $d\bar{d}$ pair (made from an unseen gluon) to form the final state.

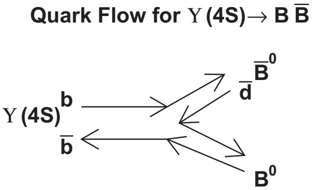


Figure 6.7: Quark flow diagram for the strong decay $\Upsilon(4S) \to B\bar{B}$. An internal $d\bar{d}$ pair combines with $\bar{b}$ and b of the parent to form the mesons in the final state.

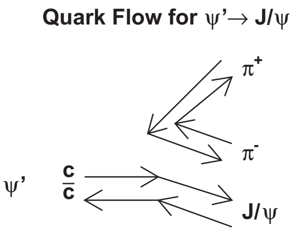


Figure 6.8: Quark flow diagram for the strong, but inhibited decay $\psi' \to J/\psi\pi^+\pi^-$. The charmonium states accommodate the $c\bar{c}$ pair, leaving the two pions by themselves. One may imagine gluons radiated by the $c\bar{c}$ pair creating the two quark pairs for the pions. This is called a disconnected diagram, with a transition rate that is suppressed.

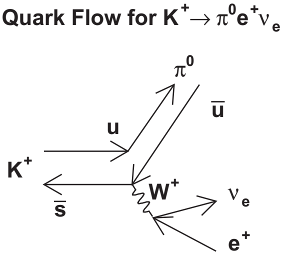


Figure 6.9: Quark flow diagram for the semi-leptonic decay $K^+ \to \pi^0 e^+ \nu_e$. Here the u quark is a spectator.

into a $D\bar{D}$ pair. This is called ‘Zweig’s rule’, after George Zweig, who first proposed that disconnected quark diagrams were inhibited. Weak decays, where a quark changes flavor, may be pictured with an internal W boson. Figure 6.9 shows the semileptonic decay $K^+ \to \pi^0 e^+ \nu_e$, where the $\bar{s}$ quark emits a W^+ and becomes a $\bar{u}$ quark. The u quark on the outside is just along for the ride, and is called a ‘spectator’. Figure 6.10 shows a bit more complicated decay $D^0 \to \bar{K}^0 \pi^+ \pi^-$.

Many examples of mesons that cannot be accommodated in a $q\bar{q}$ quark model have been found in studying the final states of B meson decay. These states are resonances that decay by strong interaction, and are characterized by line widths rather than lifetimes. A $q\bar{q}$ spin 1/2 fermion–anti-fermion pair can have quantum numbers

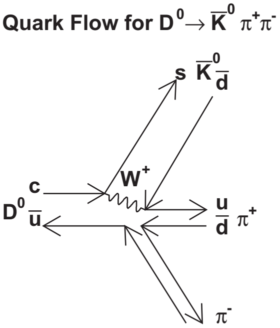


Figure 6.10: Quark flow diagram for the three-body decay $D^0 \to \bar{K}^0\pi^+\pi^-$. Here the $\bar{u}$ quark is a spectator, and there is a $d\bar{d}$ pair created to form the outgoing pions.

$$J^{\mathrm{PC}} = 0^{++}, 0^{-+}, 1^{++}, 1^{+-}, 1^{--}, 2^{++}, 2^{-+}, 2^{--}; \text{ where}$$
$$\vec{J} = \vec{L} + \vec{S}, \mathrm{P} = (-1)^{\mathrm{L+1}}, \mathrm{C} = (-1)^{\mathrm{L+S}}. \tag{6.104}$$

In the center of mass of the $q\bar{q}$ pair parity and charge conjugation are similar operations. The extra -1 that appears in the definition of parity comes from the odd intrinsic parity of a fermion–anti-fermion pair that was discussed in Chapter 2, where the positron spinors are defined. Spin enters the charge conjugation quantum number through exchange symmetry. States formed by e^-e^+ annihilation have $J^{\mathrm{PC}} = 1^{--}$. Mesons that have quantum number assignments different from this list are called "exotic". The first exotic candidate found in the final state of a B meson decay is called the $X(3872)$, with a mass of 3872 MeV and a width $\Gamma < 1.2$ MeV. It was discovered by the Belle collaboration [12] by plotting the invariant mass of $J/\psi\pi^+\pi^-$ in the final state of the decay $B^+ \to K^+J/\psi\pi^+\pi^-$. The final state particles are not exotic by themselves. In fact these particles form a principal decay mode of the second 1^{--} $c\bar{c}$ state, the ψ' at 3686 MeV, and there is a prominent peak in the mass distribution at 3686 MeV in addition to one at 3872 MeV. The dipion mass distribution favors $m_{\pi\pi} = 770$ MeV, which is the vector meson ρ (also 1^{--}) resonance, full width 150 MeV. Only the low energy side of the ρ resonance can contribute, because the sum of the two masses $3097+770 = 3867$ MeV, just 5 MeV below the peak. So there appears to be two vectors with $J^{\mathrm{PC}} = 1^{--}$ in the final state of the decay of the X. Another clue comes from the decay $X \to J/\psi\,\gamma$, that establishes the charge conjugation quantum number $C = +1$, because both the J/ψ and the γ are odd under C. The parent B^+ and the recoil K^+ both have spin zero. Analysis of the angular distribution of the final state particles by LHCb has established $J^{\mathrm{PC}} = 1^{++}$ [13]. While this assignment is accessible to $q\bar{q}$ according to the above listing, the mass does not match any $c\bar{c}$ state. In addition, the mass is almost exactly equal to the sum of masses of $D^0 + \bar{D}^{*0}$, leading to the hypothesis that it is a loosely bound four-quark molecule. Many other similar states have since been found. Perhaps one of the most convincing is the $Z(4430)^{\pm} \to \psi'\pi^{\pm}$. This object is charged, and decays into a neutral cc̄ pair plus a charged pion, so it must contain more than one $q\bar{q}$ pair. For a review of the current status of this subject, see the article by Olsen *et al.* [14].

6.7. Problems

6.1 Derive Equation (6.18)

$$\left\langle u^{\dagger(2)}(p)\right| \frac{1-\gamma^5}{2} \left| u^{(2)}(p)\right\rangle = \frac{1}{2}(1+v).$$

6.2 Check Equations (6.23) and (6.24):

$$\mathrm{Tr}(A^\dagger A) = \mathrm{Tr}(B^\dagger B) = 32m_p^2(2E_eE_\nu - p_e\cdot p_\nu)$$

$$\mathrm{Tr}(C^\dagger C) = \mathrm{Tr}(D^\dagger D) = 32m_p^2\lambda^2(2E_eE_\nu + p_e\cdot p_\nu).$$

6.3 Obtain Equation (6.28) for ^{14}O decay from Equation (6.27).

6.4 Neutron decay in the beam method. Assume thermal neutrons all with velocity 2000 m/sec and lifetime 879 s. What fraction decays in a 2-m long pipe? For a neutron current of 10^9 neutrons/s through the pipe, and a detection efficiency for the daughter protons of 10%, how many proton counts per sec do you record? Show that the proton recoil kinetic energy is about 750 eV (assumed to be zero in our analysis). The shape of the recoil proton energy distribution depends on the angular correlation parameter $a = (1-\lambda^2)/(1+3\lambda^2)$. To show this, it is necessary to consider the neutron decay as a three-body rather than a two-body final state. We lead the reader through the problem, but you are expected to confirm the calculations. First write the energy dependence of Equation (6.25) as

$$\sum_{\text{spins}} |\left\langle pe^-\bar{\nu}_e\right| T \left|n\right\rangle|^2 \sim (E_e E_\nu + a p_e E_\nu \cos(\theta_{e\nu})).$$

Now use Equation (2.67) for $2p_e E_\nu \cos(\theta_{e\nu})$:

$$2p_e E_\nu \cos(\theta_{e\nu}) = m_n^2 + m_e^2 - m_p^2 + 2E_e E_\nu - 2m_n(E_e + E_\nu),$$
$$\text{where } E_\nu = m_n - E_e - m_p - T_p = \Delta - E_e - T_p.$$

T_p is the proton kinetic energy, which as we have seen is about 10^{-6} of the proton mass, so it has to be factored out in order not to get lost. Rearranging terms and approximating $m_n^2 - m_p^2 \sim 2m_n\Delta$ gives:

$$\sum_{\text{spins}} |\left\langle pe^-\bar{\nu}_e\right| T \left|n\right\rangle|^2 \sim 2E_e(\Delta - E_e - T_p) + a(m_e^2 - 2E_e^2 + 2E_e(\Delta - T_p) + 2m_n T_p).$$

Now Equation (2.85) shows that the three-body phase space reduces to $dE_e dE_p$, so we have the recoil proton kinetic energy spectrum:

$$\frac{df(T_p)}{dT_p} = \int_{E_e^-}^{E_e^+} (2E_e(\Delta - E_e - T_p) + a(m_e^2 - 2E_e^2 + 2E_e(\Delta - T_p) + 2m_n T_p)dE_e,$$

where the limits $E_e^\pm$ come from solving the quadratic equation that follows from Equation (2.67) for $|\cos\theta_{ep}| = 1$. Show that these limits are (no approximations):

$$E_e^\pm = \frac{A(\Delta - T_p) \pm \sqrt{A^2(\Delta - T_p)^2 - (A^2 + 4m_e^2 T_p(T_p + 2m_p))(\Delta^2 - 2m_n T_p)}}{2(\Delta^2 - 2m_n T_p)},$$

where $A = \Delta^2 + m_e^2 - 2m_n T_p$. This complicated expression is just the solution to a quadratic equation that follows from the boundary curve for a three-body decay with one mass set to zero. Recall the rules from Chapter 2. For a decay $m \to m_1 + m_2 + m_3$, with $m_3 = 0$, particles 1 and 2 have simultaneous maximum energy, when particle 3 is at rest. The maximum energies follow the two-body formula: $E_1 = (m^2 + m_1^2 - m_2^2)/(2m)$. Show that this simultaneous limit is consistent with the above equation. The measurement of $f(T_p)$ is a challenge. A 750 eV proton will not go through tissue paper, so everything has to be evacuated. An accelerating voltage of 20 kV or so can be added to boost the proton energy, and then measure the distribution riding on 20 kV. The asymmetry parameter is small, so the effect on $f(T_p)$ is also modest.

6.5 Derive the $V - A$ trace theorems Equations (6.35) and (6.36).

6.6 Derive the useful identity for reducing a $V - A$ expression that includes a spin vector, Equation (6.49):

$$\begin{aligned}&(1-\gamma^5)(\not{p}_1 + m)(1 + \lambda_1\gamma^5\not{s}_1)(1+\gamma^5)\\ &\quad = (1-\gamma^5)(\not{p}_1 - \lambda_1 m\not{s}_1 + m)(1+\gamma^5).\end{aligned}$$

6.7 Prove Equation (6.53). Theorem for two-body phase space of massless particles in a three-body final state $p_1 = p_2 + p_3 + p_4$:

$$\int d_2(PS) = \int \delta^4(X - p_2 - p_3)\frac{d^3p_2}{2E_2}\frac{d^3p_3}{2E_3},$$

$$\int d_2(PS)p_2^\alpha p_3^\beta = \frac{\pi}{24}\{g^{\alpha\beta}X^2 + 2X^\alpha X^\beta\},$$

$$X^\alpha = p_1^\alpha - p_4^\alpha = p_2^\alpha + p_3^\alpha.$$

$d_2(PS)$ is Lorentz invariant, and depends only on p_2 and p_3, so the most general form must be:

$$\int d_2(PS)p_2^\alpha p_3^\beta = Ag^{\alpha\beta}X^2 + BX^\alpha X^\beta,$$

$$\text{where } A \text{ and } B \text{ are constants.} \tag{6.105}$$

Calculate the phase space integrals $\int d_2(PS)p_2 \cdot p_3$, and $\int d_2(PS)(p_2\cdot X)(p_3\cdot X)$, giving two equations to solve for A and B. Since the expressions are Lorentz invariant, you may work in the two-neutrino center of mass, where $\vec{p}_2 = -\vec{p}_3$, simplifying the integrals.

6.8 Verify the phase space integral Equation (6.67).

$$\int d^3p_1 \delta\left(m_\pi - E_1 - \sqrt{E_1^2 - m_\mu^2}\right) = \frac{4\pi p^2 E_1}{m_\pi}.$$

6.9 Verify the two-pion phase space integral Equation (6.71)

$$\int \frac{(2\pi)^4}{2m_K(2\pi)^6}\delta^4(p_k - p_1 - p_2)\frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2} = \frac{1}{16\pi m_K}\frac{p}{E}.$$

6.10 Adopt Equation (6.68) to the leptonic decay of the $K^+ \to \mu^+ + \nu_\mu$. Replace $\cos\theta_c f_\pi$ with $\sin\theta_c f_K$. Look up the relevant masses, lifetimes and branching ratios, and calculate the ratio f_K/f_π. I got 1.22.

6.11 Decays of the τ lepton. The τ, with a mass $m_\tau = 1777\,\text{MeV}$ has many decay modes, but all measurements are consistent with it being a true lepton, obeying the Dirac equation for a spin 1/2 particle. Therefore, despite the heavy mass, it may be treated just like a muon in the decay $\tau^+ \to e^+\nu_e\bar{\nu}_\tau$. The τ lifetime is 2.90×10^{-13} s, and the branching fraction to $e^+\nu_e\bar{\nu}_\tau = 17.8\%$. Check these numbers with the muon formula Equation (6.46) (I get the prediction to be about 2% low). Another interesting decay mode is $\tau^+ \to \pi^+\bar{\nu}_\tau$, with a branching fraction of 10.8%. The decay $\pi^+ \to \tau^+\nu_\tau$ is energy forbidden, but $\pi^+ \to \mu^+\nu_\mu$ is not, and the μ and τ should be interchangeable. So, it is tempting to try our analysis of pion decay on this decay mode of the τ. Write Equation (6.60) as follows:

$$\left\langle \pi^+\bar{\nu}_\tau \right| T \left| \tau^+ \right\rangle = \frac{G_F \cos(\theta_c) f_\pi}{\sqrt{2}}(p_\tau^\lambda - p_\nu^\lambda)\bar{v}(p_\tau)\gamma_\lambda(1-\gamma^5)u(p_\nu).$$

For pion decay $f_\pi \sim m_\pi$, but here the pion 4-momentum is $q \sim m_\tau$, more than a factor of ten larger. Since the q^2 dependence of f_π is unknown, we soldier on. Show that:

$$\sum_{\text{spins}} |\left\langle \pi^+\bar{\nu}_\tau \right| T \left| \tau^+ \right\rangle|^2 = \frac{G_F^2 \cos^2(\theta_c) f_\pi^2}{2} 8m_\tau^3 E_\nu.$$

Then evaluate the two-body final state phase space (remember to divide by two for the τ spin) to obtain the decay rate:

$$\Gamma(\tau^+ \to \pi^+ \bar{\nu}_\tau) = \frac{G_F^2 \cos^2(\theta_c) f_\pi^2}{4\pi} m_\tau E_\nu^2.$$

Compare this number to the experimental rate, assuming $f_\pi = m_\pi$. The numbers are surprisingly close.

6.12 Using Equations (6.75) and (6.76), show that the proton helicity for unpolarized Λ's is $\langle \sigma \cdot \hat{n} \rangle = \alpha$, where $\hat{n} = \cos\theta \hat{z} + \sin\theta \hat{x}$, is the direction of the proton momentum in the Λ rest frame, and $\alpha = 2\mathrm{SP}/(\mathrm{S}^2 + \mathrm{P}^2)$.

6.13 One can demonstrate from the machinery we have developed that a non-collinear Lorentz transformation on a massless Dirac spinor does not alter the helicity. The relevant formulas are all in Chapter 2. Begin with a massless spinor with momentum along $\hat{z}$ and $\langle \vec{\sigma} \cdot \hat{p} \rangle = +1$. Lorentz transform along $\hat{x}$. Now $p' = (\gamma E, v\gamma E, 0, E)$, and

$$\vec{\sigma} \cdot \hat{p}' = \begin{pmatrix} 1/\gamma & v \\ v & -1/\gamma \end{pmatrix}.$$

The spinor transformation matrix $S = \cosh(\chi/2) + \sinh(\chi/2)\, \gamma^0\gamma^1$, Equation (2.177) written for a transformation along $\hat{x}$. The rapidity χ is defined by $\tanh\chi = v$. Transforming the spinor gives:

$$Su(p) = \begin{pmatrix} \cosh(\chi/2) & 0 & 0 & \sinh(\chi/2) \\ 0 & \cosh(\chi/2) & \sinh(\chi/2) & 0 \\ 0 & \sinh(\chi/2) & \cosh(\chi/2) & 0 \\ \sinh(\chi/2) & 0 & 0 & \cosh(\chi/2) \end{pmatrix}$$

$$\times \sqrt{E} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \sqrt{E} \begin{pmatrix} \cosh(\chi/2) \\ \sinh(\chi/2) \\ \cosh(\chi/2) \\ \sinh(\chi/2) \end{pmatrix}.$$

The upper and lower components are the same. Show that $\vec{\sigma} \cdot \hat{p}' Su(p) = Su(p)$, an exercise in hyperbolic functions that proves that the spin follows the momentum.

6.14 $|\Delta I| = 1/2$ rule and hyperon decays was the subject of Problem 1.8 in Chapter 1. However, we have since learned that the decays of Λ and Σ hyperons involve S and P waves. From Problem 1.8 we had the three Σ decays in terms of two complex amplitudes:

$$A_0 = \frac{\sqrt{2}(A_{3/2} + A_{1/2})}{3}; \; A_+ = \frac{A_{3/2} - 2A_{1/2}}{3}; \; A_- = A_{3/2}.$$

$$\text{Giving } \sqrt{2}A_0 + A_+ - A_- = 0.$$

The subscript refers to the sign of the pion charge. In addition to the α parameters listed in Table 6.3, the γ parameters are known, but were omitted for lack of space. They are: $\gamma_0 = 0.16$; $\gamma_+ = -0.97$; and $\gamma_- = 0.98$. Attempt to construct a triangle on a two-dimensional plot with the axes labeled S wave and P wave, using the decay rates, the $|\Delta I| = 1/2$ rule formula, and our knowledge of α and γ. Nature has given us more symmetry in these three decays than is required by the $|\Delta I| = 1/2$ rule.

6.15 Adopt the analysis for $\Lambda \to pe^-\bar{\nu}_e$ to $\Sigma^- \to ne^-\bar{\nu}_e$ and compare with experiment. Use the correct V-spin SU(3) matrix element ($\langle K^0 | V_+ | \pi^- \rangle$ in Table 1.4), and $\sin\theta_c$.

6.16 Many baryons containing charm or bottom quarks have been found by experiment. The baryons with charm quarks are listed in the text: $\Lambda_c^+ = udc$, $\Sigma_c^{++} = uuc$; $\Sigma_c^+ = udc$; $\Sigma_c^0 = ddc$; $\Xi_c^+ = usc$; $\Xi_c^0 = dsc$; and $\Omega_c^0 = ssc$, all with spin/parity $(1/2)^+$ and $C = +1$. Note that the Ω_c^0 has spin 1/2, different from the one with three strange quarks, $\Omega^-(sss)$, that has spin 3/2. Make a list with quark content of the bottom quark baryons, using the same notation, adopted from the strange particles. Note that the bottom quark has charge $-1/3$, so that the charges of the bottom baryons match those of the strange baryons. The bottom quark has quantum number $B = -1$.

Charm decays to strange, which is allowed, but $b \to c$ and $s \to u$ are both forbidden, and hence have longer lifetimes. Λ_b, Σ_b, Ξ_b, and Ω_b with spin 1/2 have all been seen. Pick out a non-leptonic final state for Λ_c, Σ_c^{++}, Ξ_c^{+}, and Ω_c^{0}, and for the b quark analogs, and draw a quark flow diagram for each decay.

References

[1] G. Darius *et al.*, *Phys. Rev. Lett.* **119** (4), 042502 (2017).

[2] S. Paul, arXiv:0902.0169v2, 6 March 2009.

[3] V. Barger and R. Phillips, *Collider Physics* (Addison Wesley, New York, 1997), p. 70.

[4] *ibid*, pp. 68–72.

[5] P. A. Zyla *et al.* (Particle Data Group), *Prog. Theor. Exp. Phys.* **2020**, 083c01 (2020).

[6] J. W. Cronin and O. E. Overseth, *Phys. Rev.* **129**, 1795 (1963).

[7] C. Li and K. Schönning (BESIII Collaboration), arXiv:1902. 09098v1, 25 Feburary 2019.

[8] E. Merzbacher, *Quantum Mechanics*, 2nd edn. (John Wiley, New York, 1970), Chapter 13.

[9] N. Cabibbo, *Phys. Rev. Lett.* **10**, 531(1963).

[10] M. Aker *et al.*, *Phys. Rev. Lett.* **123**, 221802 (2019).

[11] D. M. Asner *et al.* (CLEO Collaboration), *Phys. Rev. D.* **76**, 012007 (2007).

[12] S.-K. Choi *et al.* (Belle Collaboration), *Phys. Rev. Lett.* **91**, 262001 (2003).

[13] R. Aaij *et al.* (LHCb Collaboration), *Phys. Rev. Lett.* **110**, 222001 (2013).

[14] S. L. Olsen, T. Skwarnicki, and D. Ziemenska, *Rev. Mod. Phys.* **90**, 015003 (2018).

Chapter 7

Electroweak Phenomenology

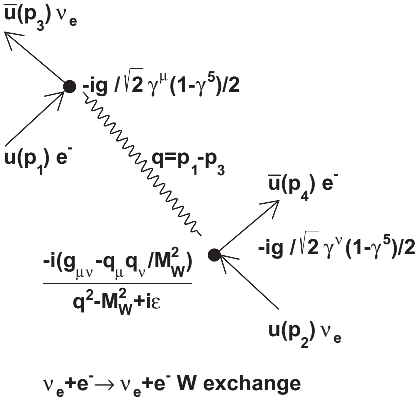
7.1. Introduction to Neutrino Scattering

The objective of this chapter is to familiarize the reader with the many aspects of electroweak theory before a discussion of the field theoretic background in the next chapter. Neutrino scattering is a good place to get started, because it was the detection of neutral currents in neutrino interactions that was one of the early predictions of the electroweak theory to be confirmed. Weak neutral currents, mediated by a massive spin one boson, called the Z, had not been observed when Weinberg published his paper with the modest title *A Model of Leptons* [1], that contained *all* of the elements of our present electroweak theory. Previous experiments, some already described in Section 1.5 of Chapter 1, had detected the quasi-elastic charged current reaction $\nu_\mu + n \to \mu^- + p$. The neutral current analog would be $\nu_\mu + n \to \nu_\mu + n$, or $\nu_\mu + p \to \nu_\mu + p$ with nothing in the final state but a recoil neutron or proton and an unobserved neutrino. Although the neutral current is about a 20% effect, it was devilishly difficult to pin it down, because of neutron background that could mimic the signal. The group at the Gargamelle heavy liquid bubble chamber at CERN succeeded in bringing the backgrounds under control and reporting the existence of weak neutral currents in 1973, 5 years after Weinberg's paper [2]. From their neutral/charged ratio an estimate could be made of the W and Z boson masses. The answer 50–100 GeV was out of reach of any fixed target accelerator at that time, and stimulated the development of proton–antiproton colliding beams to make the new fundamental particles.

Particle accelerators are not the only sources of neutrinos. The first direct demonstration of the existence of neutrinos was the pioneering experiment of Cowan and Reines [3], who detected $\bar{\nu}_e + p \to e^+ + n$ using $\sim$5 MeV antineutrinos from a reactor. As mentioned in Chapter 1, ν_e's were first detected from nuclear fusion, the power source for the sun, by Ray Davis, who counted $^{37}\mathrm{Ar}$ atoms from a tank of 378,000 L of perchloroethylene via $\nu_e + {}^{37}\mathrm{Cl} \to {}^{37}\mathrm{Ar} + e^-$. Neutrino astronomy was born when about six neutrino events were detected in a large underground water tank in coincidence with the explosion of the supernova SN1987A [4]. There is a substantial flux of cosmic ray neutrinos that comes from interactions of the primary proton cosmic radiation in the upper atmosphere. Like atmospheric neutrinos, neutrino beams at particle accelerators predominantly come from pion decay in flight. A charged pion beam is made by bombarding a metal target with accelerated protons, and then a decay drift space is created for decay in flight. The pion decay length is $c\tau = 7.8\,\mathrm{m}$, which becomes 7.80 km for a 140 GeV pion, so the drift length tends to be long. Then enough absorber (steel, concrete, earth) is placed in the beam to take out the remaining pions and the daughter muons, leaving only the neutrinos for study. This may sound quiet and peaceful, but the neutrinos themselves can be an annoying source of backgrounds. Neutrino cross-sections are very small, but an intense neutrino flux can make enough neutrons to be troublesome. The neutron background that obscured the neutral current signal in Gargamelle was made by interaction of the neutrino beam with the surrounding shielding. Since the proton beam is positively charged, positive pions are about three times more numerous than negative pions, so that the flux of ν_μ's is greater than $\bar{\nu}_\mu$'s. However, both charged current reactions $\nu_\mu + n \to \mu^- + p$, and $\bar{\nu}_\mu + p \to \mu^+ + n$ have been extensively studied in external neutrino beams. Solar and reactor electron neutrino energies are in the few MeV range, while accelerator neutrino beams that have been built range from 1 GeV to over 100 GeV. Nucleon structure plays an important role in the interpretation of neutrino–nucleon scattering. In the quark model, the muon neutrino interacts with d and $\bar{u}$ quarks, and the muon antineutrino interacts with u and $\bar{d}$ quarks, the antiquarks being in the sea. In order to apply the parton model introduced in Chapter 5 to neutrino scattering, we must derive formulas for neutrino interactions with quarks and antiquarks, and then use the parton distribution functions. We begin with neutrino electron scattering.

7.2. Neutrino Electron Scattering

We have seen that the electroweak theory groups the leptons in doublets, and charged current weak interactions give transitions within doublets with no cross coupling, Equation (6.101). The neutral current also respects the doublet structure; there is no coupling $Z \to \mu^+ e^-$, only $Z \to e^+ e^-$ or $Z \to \mu^+ \mu^-$, or $Z \to \nu_e \bar{\nu}_e$, etc. The two t channel diagrams for $\nu_e e^- \to \nu_e e^-$ are shown in Figures 7.1 and 7.2. For Figure 7.1, we have chosen p_1 and p_2 the incident four momenta for the electron and neutrino, respectively, and p_3 is the



$\bar{u}(p_3)\,\nu_e$
$-ig/\sqrt{2}\,\gamma^\mu(1-\gamma^5)/2$
$u(p_1)\,e^-$
$q=p_1-p_3$
$\bar{u}(p_4)\,e^-$
$-ig/\sqrt{2}\,\gamma^\nu(1-\gamma^5)/2$
$\frac{-i(g_{\mu\nu}-q_\mu q_\nu/M_W^2)}{q^2-M_W^2+i\varepsilon}$
$u(p_2)\,\nu_e$
$\nu_e+e^- \to \nu_e+e^-$ W exchange

Figure 7.1: Feynman diagram with spinors and propagators needed to construct an amplitude for $\nu_e + e^- \to \nu_e + e^-$ by t channel W exchange. For $q^2 \ll M_W^2$ the propagator reduces to $ig_{\mu\nu}/M_W^2$. Then identifying the Fermi coupling constant $G_F/\sqrt{2} = g^2/(8M_W^2)$ gives Equation (7.2).

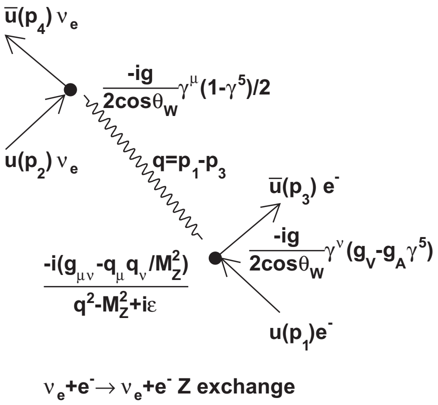

$\bar{u}(p_4)\,\nu_e$
$\frac{-ig}{2\cos\theta_W}\gamma^\mu(1-\gamma^5)/2$
$u(p_2)\,\nu_e$
$q=p_1-p_3$
$\bar{u}(p_3)\,e^-$
$\frac{-ig}{2\cos\theta_W}\gamma^\nu(g_V-g_A\gamma^5)$
$\frac{-i(g_{\mu\nu}-q_\mu q_\nu/M_Z^2)}{q^2-M_Z^2+i\varepsilon}$
$u(p_1)e^-$
$\nu_e+e^-\rightarrow\nu_e+e^-$ Z exchange

Figure 7.2: Feynman diagram with spinors and propagators needed to construct an amplitude for $\nu_e + e^- \rightarrow \nu_e + e^-$ by t channel Z exchange. The coupling of the Z to neutrinos is still $V - A$. The constant g is the same as in Figure 7.1. θ_W is a parameter of the theory, measured by experiment: $\cos\theta_W = 0.88$. The constants g_V and g_A are given for the electron in Table 7.2.

outgoing neutrino four momentum. The exchanged boson is a W^-. The Mandelstam variables are

$$\begin{aligned} p_1 + p_2 = p_3 + p_4; \quad s = (p_1 + p_2)^2; \quad t = (p_1 - p_3)^2; \\ u = (p_1 - p_4)^2. \end{aligned} \tag{7.1}$$

Note that the angle between $\vec{p}_1$ and $\vec{p}_3$ is not directly measurable, because the outgoing neutrino is unobserved, but the companion angle between $\vec{p}_2$ and $\vec{p}_4$ can be measured. The direction of the incident neutrino beam at an accelerator is known. At the W vertex a ν_e couples to an electron, and at the Z vertex a ν_e to a ν_e, and any lepton, neutrino or quark could be accommodated on the other side. While the coupling constant at the W vertex is always $(G_F/\sqrt{2})^{1/2}\gamma^\mu(1-\gamma^5)$, the coupling at the Z vertex is not always $V - A$, and depends on the charge of the lepton or quark, so it is different for e's and ν's. The g_V and g_A assignments for coupling the Z to leptons and quarks are given in Table 7.2. Antiparticles have opposite signs.

The dominant source of neutrinos at an accelerator is $\pi \rightarrow \mu + \nu_\mu$, and a ν_μ scatters only off of a μ^-. As there are no handy muon targets, it would seem that purely leptonic neutrino interactions would be not so easy to study. There are, on the other hand, plenty of target electrons around, and there are actually ν_e's and $\bar{\nu}_e$'s in the beam at a few % level, that come from K decays. So, a process like $\nu_e + e^- \to \nu_e + e^-$ can actually be observed, and a cross-section can be measured provided you can calculate the neutrino flux.

Accordingly we write down the matrix element for $e^- + \nu_e \to \nu_e + e^-$, $p_1 + p_2 = p_3 + p_4$. This t channel process with an intermediate W^- is shown in Figure 7.1:

$$\left\langle \nu_e e^- \right| T \left| \nu_e e^- \right\rangle = \frac{G_F}{\sqrt{2}} \bar{u}(p_3)\gamma^\lambda(1-\gamma^5)u(p_1)\bar{u}(p_4)\gamma_\lambda(1-\gamma^5)u(p_2). \tag{7.2}$$

Vectors p_1 and p_2 are incoming arrows, and p_3 and p_4 are outgoing arrows, so we may apply directly Equation (6.36).

$$\sum_{\text{spins}} |\left\langle \nu_e e^- \right| T \left| \nu_e e^- \right\rangle|^2 = \frac{G_F^2}{2} 256(p_1 \cdot p_2)(p_3 \cdot p_4) = 32 G_F^2 s^2,$$

$$\text{where } s = (p_1 + p_2)^2. \tag{7.3}$$

The electron mass terms do not contribute to the traces, so this formula is valid for MeV neutrino energies, relevant for the solar neutrino flux — see Problems 7.2–7.4.

Up until now we have been calculating weak interaction decay rates, but here we are interested in a cross-section. Equation (4.10) gives the general formula for a two-body to two-body scattering cross-section, and Equation (4.69) gives a convenient reduced formula for the angular distribution in the center of mass.

$$\frac{d\sigma_{i\to f}}{d\Omega} = \frac{|\langle f|\, T\, |i\rangle|^2}{64\pi^2 s} = \frac{G_F^2}{4\pi^2}s; \quad \sigma = \frac{G_F^2}{\pi}s = \frac{2G_F^2}{\pi}m_e E_\nu. \tag{7.4}$$

We have divided by two to average over the electron spins (the neutrino has only one spin state). This cross-section is proportional to s, the square of the total energy available. For neutrinos of energy $E_\nu \gg m_e$ incident on an electron at rest, $s = 2m_e E_\nu$. If converted to cm^2, this reduces to:

$$\sigma(\nu_e e^- \to \nu_e e^-) = 1.7 \times 10^{-41}\,\text{cm}^2 \times E_\nu \text{ in GeV}. \tag{7.5}$$

This is a very small cross-section. The number atoms per cm^3 in lead is $n = N_{\mathrm{AV}}\rho/MW = 0.3 \times 10^{23}/\mathrm{cm}^3$. Then the mean free path in lead is $\lambda = 1/(nZ\sigma) \sim 2 \times 10^{15}/E_\nu$ cm, which is an astronomical distance. Neutrino detectors are large, but they are not *that* large. For example, one of the early neutrino detectors at Fermilab had a 130 metric tons of instrumented steel [5]. So, you need a lot of neutrinos to have a chance of seeing anything. Low energy neutrinos pass through the earth unscathed. Reines and Cowan estimated the cross-section for $\bar{\nu}_e + p \to e^+ + n$ to be about $6 \times 10^{-44}\,\mathrm{cm}^2$, for neutrino energies in the MeV range.

Electron antineutrinos can also scatter off of electrons. Here, the process is s channel annihilation, as shown in Figure 7.3. The T matrix element for this is

$$\left\langle \bar{\nu}_e e^- \right| T \left| \bar{\nu}_e e^- \right\rangle = \frac{G_F}{\sqrt{2}} (\bar{v}(p_2)\gamma^\lambda(1-\gamma^5)u(p_1))(\bar{u}(p_4)\gamma_\lambda(1-\gamma^5)v(p_3))). \tag{7.6}$$

The antineutrino arrows are reversed, so that p_1 and p_3 are incoming, and p_2 and p_4 are outgoing. This exchanges $p_2 \to -p_3$, or $s \to t$ in Equation (7.2).

$$\sum_{\text{spins}} |\left\langle \bar{\nu}_e e^- \right| \mathrm{T} \left| \bar{\nu}_\mathrm{e} \mathrm{e}^- \right\rangle|^2 = 32 G_F^2 t^2, \tag{7.7}$$

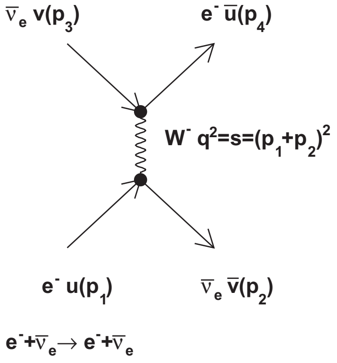

$\bar{\nu}_e$ v(p$_3$)
e$^-$ $\bar{u}$(p$_4$)
W$^-$ q^2=s=(p$_1$+p$_2$)^2
e$^-$ u(p$_1$)
$\bar{\nu}_e$ $\bar{v}$(p$_2$)
e$^-$+$\bar{\nu}_e$→ e$^-$+$\bar{\nu}_e$

Figure 7.3: Feynman annihilation diagram for $\bar{\nu}_e e^-$ scattering. This is an s channel reaction. If $s \ll M_W^2$, then the propagator is $ig_{\mu\nu}/M_W^2$, giving Equation (7.6).

where $t = (p_1 - p_3)^2 = (p_4 - p_2)^2$. In the $e^-\bar{\nu}_e$ center of mass, where all of the energies are the same, $t = -s(1 - \cos\theta)/2$, where θ is the angle between the incident antineutrino and the scattered electron. This angle can be measured if the direction of the incident antineutrino beam is known. The angular distribution is a helicity rule, because the electrons are left handed, and the antineutrinos are right handed, so the cross-section must vanish for $\theta = 0$. The differential cross-section in the center of mass is

$$\frac{d\sigma(\bar{\nu}_e e^- \to \bar{\nu}_e e^-)}{d\Omega} = \frac{G_F^2}{4\pi^4}\frac{t^2}{s}; \quad \int d\Omega \frac{d\sigma}{d\Omega} = \frac{G_F^2}{3\pi}s. \tag{7.8}$$

The antineutrino cross-section is 1/3 of the neutrino cross-section.

7.3. Neutrino Quark Scattering

The fundamental charged current neutrino reactions with nucleons are $\nu_\mu + n \to \mu^- + p$, and $\bar{\nu}_\mu + p \to \mu^+ + n$. In the quark model, the proton has valence quarks (u, u, d), and the neutron (d, d, u). The muon neutrino interacts only with the d quark: $\nu_\mu + d \to \mu^- + u$. Similarly, the muon antineutrino interacts only with the u quark: $\bar{\nu}_\mu + u \to \mu^+ + d$. The reactions are shown in Figures 7.4 and 7.5. Both are t channel W boson exchange. These processes are described by the following T matrix elements:

$$\langle \mu^- u | T | \nu_\mu d \rangle = \frac{G_F \cos\theta_c}{\sqrt{2}} \bar{u}(p_\mu)\gamma^\lambda(1 - \gamma^5)u(p_\nu)\bar{u}(p_u)\gamma_\lambda(1 - \gamma^5)u(p_d), \tag{7.9}$$

and

$$\langle \mu^+ d | T | \bar{\nu}_\mu u \rangle = \frac{G_F \cos\theta_c}{\sqrt{2}} \bar{u}(p_d)\gamma^\lambda(1 - \gamma^5)v(p_\mu)\bar{v}(p_\nu)\gamma_\lambda(1 - \gamma^5)u(p_u). \tag{7.10}$$

The Cabibbo angle θ_c was used in a similar quark formula in Chapter 6, Equation (6.4), to describe neutron decay at the quark level. $\cos\theta_c$ is replaced in Section 7.5 on quark mixing by the CKM matrix element V_{ud}. For Equation (7.9) the incoming arrows are p_ν and p_d, and the outgoing arrows are p_μ and p_u. For Equation (7.10)

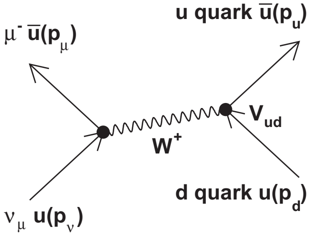


Figure 7.4: t channel reaction $\nu_\mu + d \to \mu^- + u$. V_{ud} is a CKM matrix element that replaces the Cabibbo angle $\cos\theta_c$.

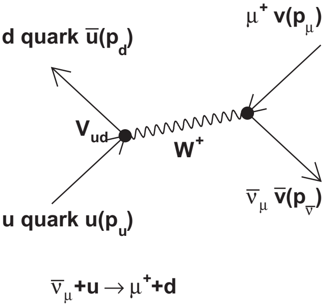


Figure 7.5: t channel reaction $\bar{\nu}_\mu + u \to \mu^+ + d$. V_{ud} is a CKM matrix element that replaces the Cabibbo angle $\cos\theta_c$.

the incoming arrows are p_μ and p_u, and the outgoing arrows are $p_{\bar{\nu}}$ and p_d. For the parton model we use the subprocess variables $\hat{s}$ and $\hat{u}$. Remember that $t = \hat{t}$. Both interactions are $V - A$, so we may apply the above theorem to write down the spin sums:

$$\begin{aligned}
\sum_{\text{spins}} |\left\langle \mu^- u \right| T \left|\nu_\mu d\right\rangle|^2 &= \frac{G_F^2 \cos^2\theta_c}{2} 256 (p_\nu \cdot p_d)(p_\mu \cdot p_u) \\
&= 32 G_F^2 \cos^2\theta_c \hat{s}^2, \\
\frac{d\sigma(\nu_\mu d \to \mu^- u)}{d\Omega} &= \frac{G_F^2 \cos^2\theta_c \hat{s}}{4\pi^2}, \\
\frac{d\sigma}{d\hat{u}} &= \frac{G_F^2 \cos^2\theta_c}{\pi}, \\
\frac{d\sigma}{dt d\hat{u}} &= \frac{1}{x}\frac{d\sigma}{dt du} = \frac{G_F^2 \cos^2\theta_c}{\pi}\delta(t + x(s+u)). \qquad (7.11)
\end{aligned}$$

This squared matrix element apart from the Cabibbo angle is the same as the one for $\nu_e e^-$ elastic scattering. After dividing by two for the u quark spin, we have used Equation (4.69) for the differential cross-section. Then we converted to invariant variables, and added a δ function at the end for insertion into the parton model. We repeat the process for the muon antineutrinos:

$$\begin{aligned}
\sum_{\text{spins}} |\left\langle \mu^+ d\right| T \left|\bar{\nu}_\mu u\right\rangle|^2 &= \frac{G_F^2 \cos^2\theta_c}{2} 256(p_\mu \cdot p_u)(p_{\bar{\nu}} \cdot p_d) \\
&= 32 G_F^2 \cos^2\theta_c \hat{u}^2, \\
\frac{d\sigma(\bar{\nu}_\mu u \to \mu^+ d)}{d\Omega} &= \frac{G_F^2 \cos^2\theta_c \hat{u}^2}{4\pi^2 s}, \\
\frac{d\sigma}{d\hat{u}} &= \frac{G_F^2 \cos^2\theta_c}{\pi} \frac{u^2}{s^2}, \\
\frac{d\sigma}{dt du} &= \frac{G_F^2 \cos^2\theta_c u^2}{\pi s^2} x\delta(t + x(s+u)). \qquad (7.12)
\end{aligned}$$

In writing down the final expressions we have assumed that the energy is high enough to ignore the muon mass. Then in terms of Mandelstam variables $(\hat{s}, t, \hat{u})$ we have $(p_\nu \cdot p_d)(p_\mu \cdot p_u) = \hat{s}^2/4$, and $(p_\mu \cdot p_u)(p_{\bar{\nu}} \cdot p_d) = \hat{u}^2/4$. If θ is the angle between the incident $\bar{\nu}_\mu$ and the outgoing μ^+ in the center of mass, then for the second reaction $\hat{u} = -(\hat{s}/2)(1 + \cos\theta)$. Equation (7.11) is isotropic in the center of mass, while Equation (7.12), for $\bar{\nu}_\mu + u \to \mu^+ + d$ vanishes for $\theta = \pi$. This is another helicity rule. Both the $\bar{\nu}_\mu$ and μ^+ are right handed, and the u and d quarks are left handed, so the initial state has $m_s = +1$ along the antineutrino direction, and the right-handed muon cannot come out backwards and conserve angular momentum.

These cross-sections look the same as the neutrino–electron cross-sections obtained above ($\cos^2\theta_c = 0.95$), yet the measured reaction in hydrogen $\bar{\nu}_\mu + p \to \mu^+ + n$ is about 10^3 times larger than Equation (7.5). How can this be, if the current quark mass is a few MeV? The answer lies in the parton model, that binds the quarks inside the proton with mass $M_p \sim 2000 \times m_e$. So the large cross-section is an example of quark binding. To apply the parton model, we need two more cross-sections for the antiquarks in the proton sea. This is achieved by exchanging the quarks between initial and final states. Just rotate the quark line about the W axis in Figure 7.4:

$$\nu_\mu + d \to \mu^- + u \text{ becomes } \nu_\mu + \bar{u} \to \mu^- + \bar{d}, \tag{7.13}$$

and

$$\bar{\nu}_\mu + u \to \mu^+ + d \text{ becomes } \bar{\nu}_\mu + \bar{d} \to \mu^+ + \bar{u}. \tag{7.14}$$

These operations interchange $\hat{s}$ and $\hat{u}$.

$$\sum_{\text{spins}} |\left\langle \mu^- \bar{d} \right| T \left| \nu_\mu \bar{u} \right\rangle|^2 = 32 G_F^2 \cos^2\theta_c \hat{u}^2;$$

$$\frac{d\sigma}{dtdu} = \frac{G_F^2 \cos^2\theta_c u^2}{\pi s^2} x\delta(t + x(s+u)), \tag{7.15}$$

and

$$\sum_{\text{spins}} |\left\langle \mu^+ \bar{u} \right| T \left| \bar{\nu}_\mu \bar{d} \right\rangle|^2 = 32 G_F^2 \cos^2\theta_c \hat{s}^2;$$

$$\frac{d\sigma}{dtdu} = \frac{G_F^2 \cos^2\theta_c}{\pi} x\delta(t + x(s+u)). \tag{7.16}$$

We now have the cross-sections to substitute into the master formula of the parton model, Equation (5.77), adapted to suit our present problem. In Chapter 5, the parton model was developed around deep inelastic *ep* scattering. We repeat Equation (5.77) here, adopted to our present analysis:

$$\frac{d\sigma(\nu_\mu + N \to \mu^- + X)}{dtdu} = \sum_i \int_0^1 dx f_i(x) \frac{d\hat{\sigma}(\nu_\mu + q_i \to \mu^- + q_j)}{dtdu}, \tag{7.17}$$

and

$$\frac{d\sigma(\bar{\nu}_\mu + N \to \mu^+ + X)}{dtdu} = \sum_i \int_0^1 dx f_i(x) \frac{d\hat{\sigma}(\bar{\nu}_\mu + q_i \to \mu^+ + q_j)}{dtdu}, \tag{7.18}$$

where N is either a neutron or a proton. We are now ready to imbed the quark level cross-sections into the parton model. Using exclusively proton PDF's, for $\nu_\mu + p \to \mu^- + X$ we have:

$$\begin{aligned}\frac{d\sigma(\nu_\mu + p \to \mu^- + X)}{dtdu} &= \frac{G_F^2 \cos^2\theta_c}{\pi}\int_0^1 dx \left(xd(x) + \frac{u^2}{s^2}x\bar{u}(x)\right)\\ &\quad \times \delta(t + x(s+u))\\ &= \frac{G_F^2 \cos^2\theta_c}{\pi}\frac{1}{s+u}\left(xd(x) + \frac{u^2}{s^2}x\bar{u}(x)\right),\\ \frac{d\sigma(\nu_\mu p \to \mu^-)}{dxdy} &= \frac{G_F^2 \cos^2\theta_c s}{\pi}(xd(x) + (1-y)^2 x\bar{u}(x)),\end{aligned} \tag{7.19}$$

where we have changed variables in the last equation. The invariant variable y was introduced in Chapter 5: $y = (u+s)/s$, or $-u/s = 1-y$.

$$\begin{aligned}\frac{d\sigma(\nu_\mu + n \to \mu^- + X)}{dtdu} &= \frac{G_F^2 \cos^2\theta_c}{\pi}\int_0^1 dx \left(xu(x) + \frac{u^2}{s^2}x\bar{d}(x)\right)\\ &\quad \times \delta(t + x(s+u)),\\ \frac{d\sigma(\nu_\mu n \to \mu^- X)}{dxdy} &= \frac{G_F^2 \cos^2\theta_c s}{\pi}(xu(x) + (1-y)^2 x\bar{d}(x)).\end{aligned} \tag{7.20}$$

On a target like ${}^{12}\mathrm{C}$, with equal numbers of protons and neutrons, these cross-sections can be combined and averaged:

$$\begin{aligned}\frac{d\sigma(\nu_\mu + {}^{12}\mathrm{C} \to \mu^- + X)}{dxdy} &= \frac{G_F^2 \cos^2\theta_c s}{2\pi}\{x(u(x) + d(x))\\ &\quad + (1-y)^2 x(\bar{u}(x) + \bar{d}(x))\},\end{aligned} \tag{7.21}$$

and

$$\begin{aligned}\frac{d\sigma(\bar{\nu}_\mu + {}^{12}\mathrm{C} \to \mu^+ + X)}{dxdy} &= \frac{G_F^2 \cos^2\theta_c s}{2\pi}\{(1-y)^2 x(u(x) + d(x))\\ &\quad + x(\bar{u}(x) + \bar{d}(x))\}.\end{aligned} \tag{7.22}$$

The proton valence quark parton distribution functions are much larger than the sea quarks, so approximately the $\nu_\mu \to \mu^-$ cross-section is isotropic in the center of mass, while the $\bar{\nu}_\mu \to \mu^+$ cross-section has the angular distribution $(1+\cos\theta)^2$. The difference is due to the fact that the quarks are all left handed, which matches the pair (ν_μ, μ^-), but is opposite to $(\bar{\nu}_\mu, \mu^+)$. The $(1-\gamma^5)'$s take care of the helicities of the leptons and quarks. The expressions in the curly brackets when integrated are of the order of unity, so the magnitudes of these cross-sections are given by the coefficients $(G_F^2 \cos^2\theta_c s)/(2\pi)$, one half of Equation (7.4). However, now $s = 2ME_\nu$, where M is the proton mass, boosting these cross-sections at constant neutrino energy by a factor of $M/(2m_e) \sim 1000$. The numerical value of the coefficient in units of E_ν in GeV is

$$\frac{G_F^2 \cos^2\theta_c M}{\pi} = 1.5 \times 10^{-38}\ \text{cm}^2/\text{GeV}. \tag{7.23}$$

This cross-section is still very small. The mean free path in iron for a 10 GeV neutrino is about 10^7 km, so in 100 m of steel the probability of an interaction is about 10^{-8}. The total mass of a neutrino detector determines its acceptance, and several hundred metric tons is a typical number. A cubic meter of iron weighs 8 metric tons. A summary paper published by the CDHS neutrino experiment at CERN gives useful numbers for a fixed target neutrino beam produced by 400 GeV protons [6]. Plots of the proton structure function $F_2(x, Q^2)$ presented in Chapter 5 are data from this experiment. The CERN SPS is a 400 GeV proton synchrotron, that produced pulses of 10^{13} protons every 14.4 s (the time required to ramp up the magnets and accelerate the injected beam). Charged pions and kaons from a beryllium target were sign selected in a wide momentum band. Only positives that decay into neutrinos, or only negatives that decay into antineutrinos were transported into a 292-m long evacuated decay tunnel. After the tunnel 464-m of earth and iron shielded the detector from charged particle background — mostly muons that accompanied the neutrinos. The detector was an array of 21 instrumented magnetized iron toroids, 3.75-m in diameter and 0.5-m thick, with an average field of 1.65 T, arranged perpendicular to the neutrino beam direction.

The first 10 toroids served as the target, and the downstream magnetized iron measured the momentum of the muon made by the charged current reaction $\nu_\mu + d \to \mu^- + u$ or $\bar{\nu}_\mu + u \to \mu^+ + d$. The sign of the magnetic field was chosen to focus the muons. The total target was about 900 tons, with a fiducial mass of 250 tons. The data sample of 640,000 neutrinos was extracted from 4.3×10^6 triggers for a total run of 1.3×10^{18} protons on target, about 0.2 coulombs of charge. The neutrino energy range was $20\,\mathrm{GeV} \leq E_\nu \leq 200\,\mathrm{GeV}$, with $\langle E_\nu \rangle = 50\,\mathrm{GeV}$. The neutrino energy spectrum decreased exponentially, but this was compensated by the linear increase in the cross-section. The total neutrino energy was defined as the sum of the muon energy measured in the toroids plus the hadronic energy obtained from the instrumented target, that served as a calorimeter. The total cross-sections used to normalize the broad band flux were $\sigma(\nu_\mu + N \to \mu^- + X)/E_\nu = 0.686 \pm 0.019 \times 10^{-38}$ cm^2/GeV, and $\sigma(\bar{\nu}_\mu + N \to \mu^+ + X)/E_{\bar{\nu}} = 0.339 \pm 0.01 \times 10^{-38}$ cm^2/GeV, about a factor of two smaller [7]. N refers to a nucleon — the cross-section for a target with equal numbers of protons and neutrons, like Equations (7.21) and (7.22). The structure functions were corrected for the fact that iron is about 10% neutron rich. The cross-sections are smaller than Equation (7.23). Assuming $x(u(x) + d(x))$ is dominant, the ratio of Equations (7.21) to (7.22) is three to one. The total cross-sections were measured in a narrow band beam, where the spatial and energy distributions of the neutrinos were more sharply defined. The detector was the same, but the configuration used to define the pion/kaon beam entering the decay pipe was different. The secondary charged beam of pions and kaons was magnetically separated into a specific momentum, and its beam intensity was monitored, as were the muons from decay in flight that accompanied the neutrinos. The result of this calibration is shown in Figure 7.6.

A tagged neutrino beam, in which the flavor of the neutrino is determined by observation of its source, for example a fully reconstructed $K^+ \to \pi^0 e^+ \nu_e$ gives electron neutrinos of known four momentum, requires some 10^8 events on the upstream side of a shielding wall for one detected neutrino, and that is a challenging task for present day instrumentation. Neutrino flavor mixing,

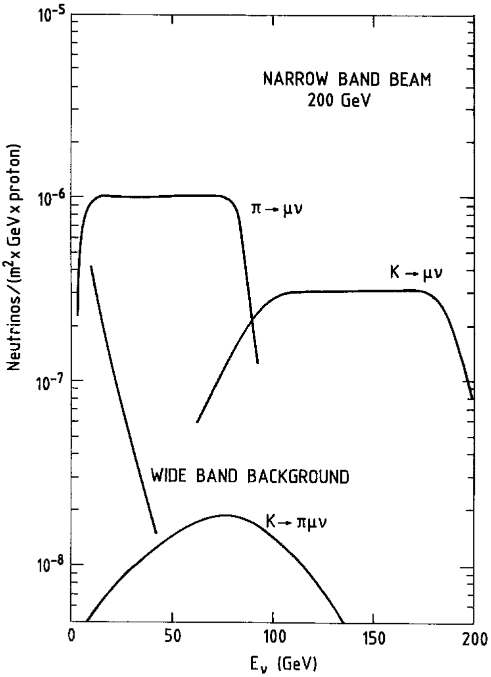


Figure 7.6: Absolute calibration of the neutrino flux from a 200-GeV charged pion/kaon beam. The flux from pion decay in flight is the largest, but cuts off around 100 GeV, where the $K \rightarrow \mu\nu$ picks up. The lower curve is from three body K decays. The wide band flux is a background. The y-axis is neutrinos/m^2 · GeV · proton, so for example 10^{13} protons on target gives $10^7 \nu_\mu/\text{m}^2 \cdot \text{GeV}$ from pion decay. The area is measured at the detector, some 750 m from the target. These flux measurements gave the total cross-sections in the text. Credit [7].

a phenomenon recently discovered, motivates the tagging of flavors at the source, to compare with the flavor at the interaction point. The reconstructed neutrino line of flight and energy are useful parameters to combine with the timing to reduce accidental coincidences.

7.4. $W^\pm$ Intermediate Vector Bosons

The charged intermediate vector bosons $W^\pm$ have appeared in our diagrams, but not in our formulas up to this point, although a T matrix element involving the W was presented in Equation (6.4).

The reason for the approximation that replaces $g^2/(8M_W^2) \to G_F/\sqrt{2}$ (Equation (6.5)) is that the energy involved in the transition is small compared to M_W. The cross-section is proportional to s, and grows without limit as the neutrino energy increases. Cross-sections in electrodynamics that we derived in Chapter 4 do not have this property. Most of them decrease as ∼1/s. This divergence comes from the point interaction and the absence of a propagator. It is possible to create a real 'on mass shell' W boson by s channel annihilation: $u\bar{d} \to W^+$. This is an example of the Drell–Yan process in the parton model, introduced in Chapter 5. The proton carries the u quark, and the antiproton the $\bar{d}$ quark, with momentum fractions x_1 and x_2, specified by proton parton distribution functions. If a real W is produced $\hat{s} = x_1 x_2 s = M_W^2$; $i\epsilon$ is then needed in the denominator of the W propagator to preserve convergence, and is directly related to the line width of the W: $\epsilon = M_W \Gamma_W$. This process is shown in Figure 7.7, and has the T matrix element:

$$\begin{aligned}
-i \left\langle \mu^+ \nu_\mu \right| T \left| u\bar{d} \right\rangle &= \frac{g^2 \cos\theta_c}{8} \bar{v}(p_2)\gamma^\mu(1-\gamma^5)u(p_1) \\
&\quad \times \left\{ \frac{i(-g_{\mu\nu} + (p_1+p_2)_\mu (p_3+p_4)_\nu/(M_W^2)}{\hat{s} - M_W^2 + iM_W\Gamma_W} \right\} \\
&\quad \times \bar{u}(p_3)\gamma^\nu(1-\gamma^5)v(p_4). \qquad (7.24)
\end{aligned}$$

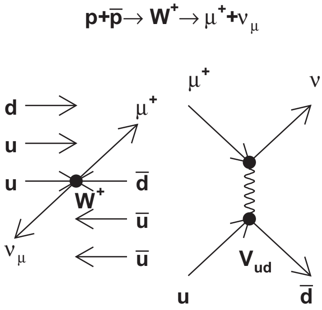


Figure 7.7: Parton model diagram for $p\bar{p} \to W^+$. Protons are incident from the left, and antiprotons from the right. The proton has two spectator quarks, d and u, and the antiproton two spectator antiquarks $\bar{u}$ and $\bar{u}$. None of the quarks carry transverse momentum. Feynman diagram for the quark reaction is shown on the right.

The four momenta are $p_1(u \text{ quark}) + p_2(\bar{d} \text{ quark}) = p_3(\nu_\mu) + p_4(\mu^+)$. The Mandelstam variables are $\hat{s} = (p_1 + p_2)^2$, $\hat{t} = (p_1 - p_3)^2$, and $\hat{u} = (p_1 - p_4)^2$. In the parton model there is no transverse momentum. The initial momenta $\vec{p}_1$ and $\vec{p}_2$ are collinear, and cancel each other in the center of mass. All particles except the W^+ are assumed massless, and the center of mass coincides with the W rest frame. The first parts of the W propagator in numerator and denominator, $-ig_{\mu\nu}/\hat{s}$ are the same as that for the photon, and we argued that the numerator represented a sum over polarization states, at least when sandwiched between conserved electromagnetic currents. The numerator in the curly brackets of Equation (7.24) is also a sum over the polarization states of the massive vector boson. Both vector bosons have spin one, but the photon has no mass, so it has only two spin states. The number of spin states changes from three to two discontinuously as the mass goes to zero. The second part of the polarization sum $(p_1 + p_2)_\mu (p_3 + p_4)_\nu / M_W^2$ is the W four momentum $P_{W\mu} P_{W\nu}/(M_W^2)$ written to facilitate the application of the Dirac equation, that cancels this term. In many, but not all, applications of the propagator this term does not contribute to the trace.

The W has a rest frame, and the description of the spin four vector is similar to the one used to describe electron spin in Section 4.10 of Chapter 4. The Dirac equation has the huge advantage of explicitly containing the γ matrices, so electron spin operators are built into the Lagrangian, a feature not shared by the vector bosons. There are three components in the W rest frame, and for motion along the z axis the transverse components are invariant. Thus,

$$\begin{aligned} p_W &= (E, 0, 0, |\vec{p}|), \quad \epsilon^{\lambda=\pm 1} = \mp(0, 1, \pm i, 0)/\sqrt{2}, \\ \epsilon^{\lambda=0} &= (|\vec{p}|, 0, 0, E)/M_W. \end{aligned} \tag{7.25}$$

In the W rest frame, the polarization states $\epsilon^{\lambda=\pm 1}$ are the same, and $\epsilon^{\lambda=0} = (0, 0, 0, 1)$, with $p_W = (M_W, 0, 0, 0)$. The signs for $\epsilon^{\lambda=\pm 1}$ are chosen to match the $Y_1^{\pm 1}$ spherical harmonics, and represent right and left circular polarization. As wave functions of the angular momentum operator, $\lambda = \pm 1$ correspond to $|S = 1, S_z = \pm 1\rangle$, and $\lambda = 0$ corresponds to $|S = 1, S_z = 0\rangle$. The photon has no component $\epsilon^{\lambda=0}$. The four vector $\epsilon^{\lambda=0}$ in Equation (7.25) has the same form as a positive helicity electron spin (Equation (4.107) in Section 4.10), but for a spin one vector it represents helicity zero. This may seem confusing, but it works, as we shall soon demonstrate. All three components are normalized space-like vectors, are orthogonal to each other, and also to the W four momentum in any frame:

$$\epsilon^{(\lambda)\mu}\epsilon_{\mu}^{\sigma *} = -\delta_{\lambda,\sigma},\ \epsilon_{\mu}^{\lambda} p_W^{\mu} = 0, \text{ all three components of } \lambda. \tag{7.26}$$

With this introduction it is possible to make sense of the claim that:

$$\sum_{\lambda} \epsilon_{\mu}^{\lambda *}\epsilon_{\nu}^{\lambda} = -g_{\mu\nu} + \frac{p_{\mu}p_{\nu}}{M_W^2}. \tag{7.27}$$

Dot both sides into p^{ν}. The left-hand side vanishes, and so does the right-hand side, because $-g_{\mu\nu}p^{\nu} + p_{\mu} = 0$. If you dot both sides into $\epsilon^{(\sigma)*\nu}$, where the superscript σ refers to a polarization state, and ν to a four vector component, you again obtain a valid equation:

$$\sum_{\lambda} \epsilon_{\mu}^{\lambda *}\epsilon_{\nu}^{\lambda}\epsilon^{(\sigma)*\nu} = -\sum_{\lambda} \epsilon_{\mu}^{\lambda *}\delta_{\lambda,\sigma} = -\epsilon_{\mu}^{\sigma *}. \tag{7.28}$$

The right-hand side gives the same thing, because $\epsilon \cdot p = 0$. Hence Equation (7.27) is a completeness relation for a spin one massive boson.

For a photon, we could make a gauge transformation on the polarization vector $\epsilon_{\mu} \rightarrow \epsilon_{\mu} + ak_{\mu}$ without any effect on the physics, because $k \cdot k = 0$. For the W, this is no longer true, $k \cdot k = M_W^2$. It is of interest to look at this problem more closely. Appeal to Equation (3.103) for the electromagnetic Lagrangian density:

$$\mathcal{L} = -F_{\mu\nu}F^{\mu\nu}/4 = -(\partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu})(\partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu})/4. \tag{7.29}$$

The field equations are

$$-\Box A^{\nu} + \partial^{\nu}(\partial_{\mu}A^{\mu}) = 0. \tag{7.30}$$

Choice of the gauge condition $\partial_{\mu}A^{\mu} = 0$ gives the wave equation for the vector potential $\Box A^{\nu} = 0$. In analogy with the Klein–Gordon equation, we could add a mass term to the electromagnetic

Lagrangian $-m^2 A^\nu A_\nu/2$. Then Equation (7.30) acquires an extra term:

$$-\Box A^\nu + \partial^\nu(\partial_\mu A^\mu) + m^2 A^\nu = 0. \tag{7.31}$$

Now taking the partial derivative ∂_ν gives

$$-\Box\partial_\nu A^\nu + \Box\partial_\mu A^\mu + m^2\partial_\nu A^\nu = 0. \tag{7.32}$$

The first two terms cancel, leaving the mass term to require $\partial_\nu A^\nu = 0$. It is no longer a gauge condition to be freely chosen.

Hence the presence of a mass term is incompatible with gauge invariance. The electroweak theory is a unification of electromagnetism and the weak interaction, and since gauge invariance is an important ingredient of electromagnetism, it is natural to extend the concept to the electroweak theory. The W boson, however, is not massless; in fact its mass is not small — 80 GeV. So, it seems that we have a dilemma. In order to accommodate gauge invariance the theory begins with all four bosons (three W's and a B) having zero mass. The quarks and leptons are also massless. Mass is generated by a process called 'spontaneous symmetry breaking', that creates the celebrated Higgs boson, and gives mass to three of the four initial bosons, now called $W^\pm$ and Z, with the fourth, the vector potential A, remaining massless. How this works in field theory will be described in Chapter 8. For now, we accept that it can be done, and gain some experience in using the results.

The Tevatron at Fermilab was a pbar-p collider that operated at $\sqrt{s} = 2$ TeV. It ran for about 25 years, ceasing operation in 2011. In order to calculate a cross-section for $W^\pm$ production in pbar-p collisions, we need to evaluate the sub-cross-section from Equation (7.24), and imbed it in a pair of parton distribution functions. The spin configuration is shown in Figure 7.8. Choosing the $+z$ axis along the proton beam, the initial state has $S_{1z} = -1/2$ for the u quark, and also $S_{2z} = -1/2$ for the $\bar{d}$ quark, because they have opposite helicities and opposite momenta. So, $S_z = -1$ is the spin of the W^+. The production tends to follow the flow of charge, that is W^+ tend to go in the proton direction, or, if you prefer, the u quarks in the proton carry a larger fraction of the momentum than the d quarks.

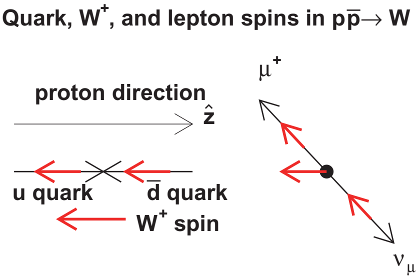


Figure 7.8: Spin diagrams for $u + \bar{d} \to W^+$, and the subsequent decay $W^+ \to \mu^+\nu_\mu$. As a result of the polarization, the positive muon prefers to go opposite to the positive proton beam.

So, the W^+ tends to go opposite to its spin direction. After production the W^+ decays: $W^+ \to \mu^+ + \nu_\mu$. The μ^+ is right handed, so it prefers to go along $-\hat{z}$, opposite to the W^+ momentum. In the reaction quark $\to W \to$ lepton the final state lepton follows the initial state quark, and in this case the final state lepton is a neutrino. The charge asymmetry in muons or electrons from $W^\pm$ decay at the Tevatron comes from the interplay of these two opposite effects. The longitudinal momentum of the W is not directly observed — only the momentum of the daughter muon. The W has spin one, and this makes a big difference between $W^+ \to \mu^+ + \nu_\mu$ and $\pi^+ \to \mu^+ + \nu_\mu$. The positive muon in pion decay has to be left handed.

The quark level cross-section can be evaluated using the techniques that we have developed. The spin sum of $|\langle f| T |i\rangle|^2$ is

$$\sum_{\text{spins}} |\langle \mu^+\nu_\mu| T |u\bar{d}\rangle|^2 = \frac{G_F^2 M_W^4 \cos^2(\theta_c)}{2((\hat{s} - M_W^2)^2 + M_W^2\Gamma_W^2)} \times \mathrm{Tr}\{\not{p}_2\gamma^\mu(1-\gamma^5)\not{p}_1\gamma^\nu(1-\gamma^5)\} \times \mathrm{Tr}\{\not{p}_3\gamma_\mu(1-\gamma^5)\not{p}_4\gamma_\nu(1-\gamma^5)\}. \quad (7.33)$$

Here we have substituted G_F for g^2. Equation (6.36) can be directly applied to this expression. The incoming arrows are p_1 and p_4, and the outgoing arrows are p_2 and p_3, so the contraction of the two tensors from the traces gives:

$$\begin{aligned}&\mathrm{Tr}\{\not{p}_2\gamma^\mu(1-\gamma^5)\not{p}_1\gamma^\nu(1-\gamma^5\}\mathrm{Tr}\{\not{p}_3\gamma_\mu(1-\gamma^5)\not{p}_4\gamma_\nu(1-\gamma^5)\}\\ &\quad = 256(p_1\cdot p_4)(p_2\cdot p_3) = 64\hat{u}^2. \end{aligned} \tag{7.34}$$

The final answer for the square of the transition matrix element is

$$\sum_{\text{spins}} |\left\langle \mu^+\nu_\mu\right| T \left|u\bar{d}\right\rangle|^2 = \frac{32G_F^2 M_W^4\cos^2(\theta_c)\hat{u}^2}{(\hat{s}-M_W^2)^2 + M_W^2\Gamma_W^2}. \tag{7.35}$$

In the W rest frame, the particle energies are all equal: $E_1 = E_2 = E_3 = E_4 = E = \sqrt{\hat{s}}/2 = M_W/2$; and $\hat{u} = -M_W^2(1-\cos\theta_{14})/2 = -M_W^2(1+\cos\theta_{13})/2$ that confirms the assertion that the lepton — a muon neutrino — follows the quark. Inside the proton are both right- and left-handed quarks, so we must divide by four for the spin average. Equation (7.41) then becomes:

$$\frac{1}{4}\sum_{\text{spins}} |\left\langle \mu^+\nu_\mu\right| T \left|u\bar{d}\right\rangle|^2 = \frac{2G_F^2 M_W^6\cos^2(\theta_c)(1+\cos\theta_{13})^2}{\Gamma_W^2}. \tag{7.36}$$

Equation (4.69) can be used to get the parton cross-section:

$$\frac{d\hat{\sigma}}{d\hat{\Omega}} = \frac{G_F^2\cos^2(\theta_c)M_W^4(1+\cos\theta_{13})^2}{32\pi^2\Gamma_W^2}, \tag{7.37}$$

and

$$\hat{\sigma}(u\bar{d}\to W^+\to\mu^+\nu_\mu) = \frac{G_F^2\cos^2(\theta_c)M_W^4}{6\pi\Gamma_W^2}. \tag{7.38}$$

We will add up all of the parts to the W line width presently. For now, we take the world averages: $M_W = 80.4\,\mathrm{GeV}$; $\Gamma_W = 2.06\,\mathrm{GeV}$; $G_F = 1.166\times 10^{-5}\,\mathrm{GeV}^{-2}$; $\theta_c = 13°$; and finally $1\,\mathrm{GeV}^{-2} = 0.389\times 10^{-27}\,\mathrm{cm}^2$. The arithmetic gives a parton cross-section for $W\to\mu\nu$:

$$\hat{\sigma}(u\bar{d}\to W^+\to\mu^+\nu_\mu) = 13\times 10^{-33}\,\mathrm{cm}^2 = 13\,\mathrm{nb}. \tag{7.39}$$

Figure 7.8 only shows the annihilating quarks, without the rest of the proton and antiproton that in the parton model are spectators to the event (see Figure 7.7). In fact, the W's are produced with small transverse momentum, and the remaining hadronic debris, called the 'underlying event' tends to be rather quiet, giving support to the parton picture. Gluons play no role directly in the Drell–Yan process, but gluon radiation can lead to transverse momentum. Measurement of final state jet activity shows that 20% of all W events had at least one jet with $E_T > 15\,\mathrm{GeV}$, where E_T is the jet transverse energy, defined as $E_T = E \sin\theta$ [8]. This number is a measure of QCD radiative corrections to the parton model. Recall the Drell–Yan formula for the mass $M_W^2 = x_1 x_2 s$, where x_1 is the u quark momentum fraction in the proton and x_2 is the $\bar{d}$ quark momentum fraction in the antiproton. With $M_W^2 = 6400\ \mathrm{GeV}^2$, and $s = 4 \times 10^6\ \mathrm{GeV}^2$, the two $x's \sim 0.04$, well within reach of the parton distribution functions. The momentum fraction carried by the W is $x_1 - x_2$. The inclusive cross-section for $p\bar{p} \to W \to e\nu$ at 1.8 TeV [8] is

$$\sigma(p\bar{p} \to W \to e\nu) = 2490 \pm 120 \times 10^{-36}\ \mathrm{cm}^2 = 2.5\ \mathrm{nb}. \tag{7.40}$$

Hence the ppbar cross-section is smaller than the quark cross-section: $\sigma(p\bar{p}) \sim 0.2 \times \hat{\sigma}(u\bar{d})$. However, this is still a quite respectable cross-section, about 10^5 times larger than the neutrino cross-sections obtained above. At a luminosity $L = 2 \times 10^{32}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$, this cross-section gives 0.5 events/s for single W production and leptonic decay. The experimental signature is a high p_T charged lepton (e or μ), and an unseen neutrino that balances p_T. W bosons have no strong interactions, but W production is only 'half weak'. The production vertex is $\sim \sqrt{G_F}$ and the subsequent decay happens 100% of the time, and costs you nothing. You only have to pay for the branching ratio, that for $\mu\nu$ is 10%.

The parton model expression for the ppbar cross-section is

$$\begin{aligned}\sigma(p\bar{p} \to W^+ \to \mu^+\nu_\mu) = \frac{K}{3}\int_0^1 dx_1 \int_0^1 dx_2 (u(x_1)d(x_2)\\ + \bar{d}(x_1)\bar{u}(x_2))\hat{\sigma}(x_1, x_2).\end{aligned} \tag{7.41}$$

The factor of 1/3 is a color factor. There are nine pairs of colors for $u\bar{d}$, and three of them are neutral, required for production of the colorless W. 'K' is called a K factor, approximately $K \sim 1.2$. The K factor is needed to match the experimental cross-section. When you multiply two of them together the normalization of the parton distributions obtained from deep inelastic lepton or neutrino scattering, as described in Chapter 5, underestimates the Drell–Yan cross-section by about 20%. This effect can be explained with perturbative QCD. In Equation (7.41), the parton distributions are all for the proton, and the second term allows the $\bar{d}$ quark to be in the proton sea. The cross-section $\hat{\sigma}(x_1, x_2)$ includes the relativistic Breit–Wigner (BW) in Equation (7.35). Integration over the angles and dividing by four for the initial spin average gives the total cross-section with the BW denominator:

$$\frac{1}{4}\sum_{\text{spins}}\int d\Omega |\langle f|\,T\,|i\rangle|^2 = \frac{32\pi G_F^2 M_W^4 \cos^2(\theta_c) x_1 x_2 s}{3((\hat{s}-M_W^2)^2 + M_W^2\Gamma_W^2)}. \tag{7.42}$$

The BW can be treated approximately as a δ function, for $\Gamma_W \ll M_W$. The integral is elementary. Since

$$\int_{-\infty}^{+\infty} \frac{dx}{x^2+a^2} = \frac{\pi}{a}.$$

So, approximately

$$\frac{1}{(\hat{s}-M_W^2)^2 + M_W^2\Gamma_W^2} = \frac{\pi}{M_W\Gamma_W}\delta(\hat{s}-M_W^2). \tag{7.43}$$

Then Equation (7.42) can be written:

$$\frac{1}{4}\sum_{\text{spins}}\int d\Omega |\langle f|\,T\,|i\rangle|^2 = \frac{32\pi^2 G_F^2 M_W^3 \cos^2\theta_c}{3\Gamma_W} x_1 \delta(x_1 - (M_W^2)/(x_2 s)), \tag{7.44}$$

where $x_2 s$ has been canceled by rearranging the δ function. The momentum fractions x_1 and x_2 are coupled together by $x_1 x_2 s = M_W^2$. Neither one can be zero. If you integrate first over dx_1 using the delta function, the lower limit on $x_2 = M_W^2/s = 0.0016$ at the Tevatron. Equation (7.41) becomes

$$\begin{aligned}\sigma(p\bar{p} \to W^+ \to \mu^+\nu_\mu) = {} & \frac{32\pi^2 G_F^2 \cos^2(\theta_c) M_W^5 K}{9s\Gamma_W} \\ & \times \int_{M_W^2/s}^{1} \frac{dx_2}{x_2}(u(M_W^2/sx_2)d(x_2) \\ & + \bar{d}(M_W^2/sx_2)\bar{u}(x_2)).\end{aligned} \tag{7.45}$$

The integration looks daunting, but it is easy to do with a computer. The parton distribution functions depend on x and Q^2, and they are available in look-up tables, for example in the CERN library of computer codes. So, set $Q^2 = M_W^2$, pick $M_W^2/s \leq x_2 \leq 1$, look up the PDF's, and sum them. The constant in front of the integral is 0.86 nb for $s = 2\,\mathrm{TeV}$, so the integral must be about 2.5 — check it out!

The mass of the $W^\pm$ bosons has been measured at the Tevatron by a technique known as the 'transverse mass'. The leptonic $W^\pm$ two-body decays involve an unseen neutrino. Decays into quarks $W^+ \to u\bar{d}$ give two charged quarks in the final state, but we are not permitted to measure their energies directly, and as we have mentioned earlier, the experimental jet energy resolution is not good enough to give an accurate value for the W mass. Additionally the QCD di-jet background in the 80-GeV mass range exceeds the W signal by about a factor of 100. An important component of any collider detector is 'hermeticity', that is the ability to detect all of the visible transverse energy from a given event. We measure transverse energy because a large amount of longitudinal energy remains unseen in the beam pipe. Hermeticity serves as a neutrino detector, because neutrinos carry away transverse energy, that leads to a E_T imbalance. The high p_T charged electron or muon is balanced by the unseen neutrino, and the missing E_T, written as $\not{E}_T$, is defined as

$$\vec{\not{E}}_T = -\vec{p}_T^{\,l} - \sum_i \vec{E}_{Ti}. \tag{7.46}$$

The transverse vector $\vec{\not{E}}_T$ is by hypothesis the $\vec{p}_T$ of the neutrino. If there are two or more neutrinos in the event, there is still only one $\not{E}_T$. The sum is over all of the directed transverse energy detected by the calorimetry. The fact that this technique works is a confirmation of the assumptions behind the parton model, discussed in Chapter 5. At the Tevatron the energy available, 2 TeV, is 25 times M_W, but W production events are relatively quiet, without, say, an accompanying 800 GeV of jet E_T. That sort of transverse energy simply does not appear. As quoted above, 80% of all W candidate events have less than 15 GeV E_T jet activity. In order to take out any smearing of

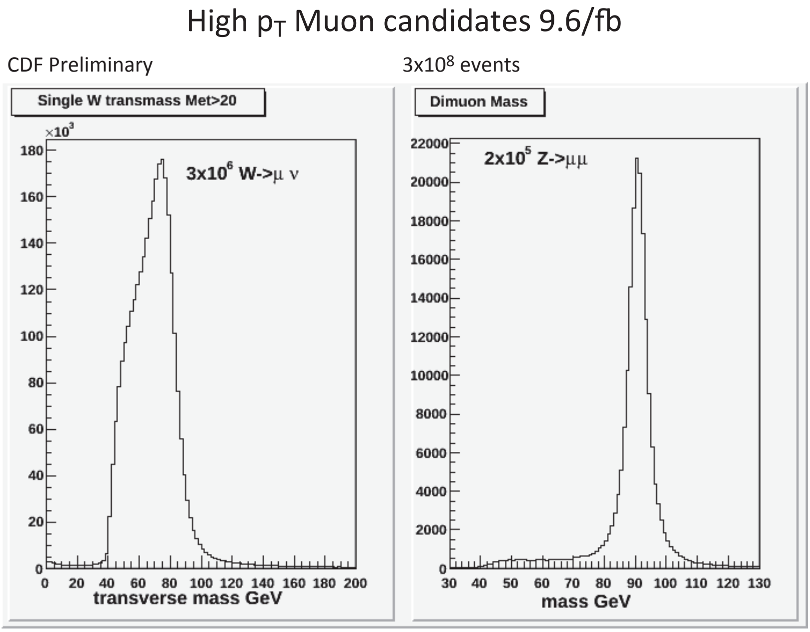


Figure 7.9: Tevatron data for the W transverse mass distribution, compared to $Z \to \mu^+\mu^-$, where a true invariant mass is reconstructed. There are 3,000,000 $W \to \mu\nu$ candidates and 200,000 $Z \to \mu^+\mu^-$ candidates in the complete data sample. The integrated luminosity is $9.6 \times 10^{39}\mathrm{cm}^{-2}$. Credit CDF Collaboration.

the charged lepton $\vec{p}_T$, the mass M_W is inferred from the transverse mass defined as

$$M_T^2 = 2(p_T^l p_T^\nu - \vec{p}_T^{\,l} \cdot \vec{p}_T^{\,\nu}), \tag{7.47}$$

where $\vec{p}_T^{\,\nu} = \vec{\not{E}}_T$ defined in Equation (7.46). Figure 7.9 shows data from the Tevatron at $\sqrt{s} = 2\,\mathrm{TeV}$. The plot with 3×10^6 $W \to \mu + \nu$ candidates for an integrated luminosity of $L = 9.6 \times 10^{39}$ cm^{-2} corresponds, given the cross-section quoted above, to an overall detection efficiency of about 12%. Z production cross-sections are the subject of Section 7.6, but the overall Z detection efficiency is slightly less than the W, because two charged leptons have to be seen. The high mass edge of the transverse mass plot is sensitive to M_W. Note that the true mass, 80 GeV, is near the peak of the distribution, not the half way point. M_Z is measured directly from the usual formula:

$$M_Z^2 = 2(E^+E^- - \vec{p}^{\,+} \cdot \vec{p}^{\,-}), \tag{7.48}$$

where these are total charged lepton momenta, not just 'transverse'. These plots are very clean, with minimal background. The full width at half maximum on the M_Z plot is about 8 GeV, dominated by momentum resolution for the muons. The resolution width is known, so that the width Γ_Z can be obtained from this type of plot. The width Γ_W can also be obtained from the transverse mass plot, but not from the shape of the edge. Rather the fact that the Breit–Wigner falls off more slowly than a Gaussian which characterizes the measurement errors leads to a measurement of Γ_W from the high transverse mass tail of the distribution. Both M_W and Γ_W have been directly measured at LEPII at CERN, where the cross-section for W pair production in e^-e^+ collisions near pair threshold, 161 GeV, is sensitive to both parameters. LEPII also furnished us with precise measurements of W branching ratios. Equation (7.40) from the Tevatron measures the product of the production cross-section times the branching ratio, but you do not know how to separate the two. All you can learn is the ratios of various branching ratios, like $\Gamma(W \to \mu\nu)/\Gamma(W \to e\nu)$. A $W^\pm$ pair, on the other hand, allows one to measure the individual branching ratios, because you know that you have a W. Top quark pairs are one source of W pairs. This does not give the partial widths or the total width in GeV, only the branching ratios.

We now return to the question of the polarization vector for the W boson. Equation (7.30) for the reaction $u\bar{d} \to W^+ \to \mu^+\nu_\mu$ can be split into two parts: $u\bar{d} \to W^+$ and $W^+ \to \mu^+\nu_\mu$. We are interested in probing the production part: $u\bar{d} \to W^+$ to learn the W^+ polarization. We think we know what the answer is, because everything is collinear, and with $\hat{z}$ in the proton direction the initial spin has $m_z = -1$. We may drop the BW, and just analyze the matrix element:

$$-i\left\langle W^+\right| T \left|u\bar{d}\right\rangle = \frac{-ig\cos\theta_c}{2\sqrt{2}}\bar{v}(p_2)\not{\epsilon}^{\lambda *}(1-\gamma^5)u(p_1), \tag{7.49}$$

where ϵ^λ are the polarization vectors defined by Equation (7.25), and only $\lambda = -1$ should contribute. The evaluation of the traces is straightforward, involving terms that we have seen before.

$$\sum_{\text{spins}} |\left\langle W^+ \right| T \left| u\bar{d} \right\rangle|^2 = g^2 \cos^2\theta_c \sum_\lambda \left\{ g^{\alpha\mu} g^{\beta\nu} + g^{\alpha\nu} g^{\beta\mu} - g^{\alpha\beta} g^{\mu\nu} \right.$$
$$\left. + i\epsilon^{\alpha\mu\beta\nu} \right\} \epsilon_\alpha^{\lambda *} p_{1\mu} \epsilon_\beta^\lambda p_{2\nu}. \tag{7.50}$$

In the W rest frame the lower index four vectors needed for the $\epsilon^{\alpha\mu\beta\nu}$ term (remember $\epsilon^{0123} = -1$) are

$$\begin{aligned} p_1 &= (E, 0, 0, -E), \quad p_2 = (E, 0, 0, E), \quad \epsilon^0 = (0, 0, 0, -1), \\ \epsilon^{+1} &= (0, 1, i, 0)/\sqrt{2}, \quad \epsilon^{+1*} = (0, 1, -i, 0)/\sqrt{2}, \\ \epsilon^{-1} &= (0, -1, i, 0)/\sqrt{2}, \quad \epsilon^{-1*} = (0, -1, -i, 0)/\sqrt{2}, \\ E &= M_W/2. \end{aligned} \tag{7.51}$$

It is straightforward to work with the four vector dot products for each value of $\lambda = \pm 1, 0$, but handling the $\epsilon^{\alpha\mu\beta\nu}$ term requires some care:

$$\sum_\lambda \{ \epsilon^\lambda \cdot p_1 \epsilon^{\lambda *} \cdot p_2 + \epsilon^\lambda \cdot p_2 \epsilon^{\lambda *} \cdot p_1 + 2E^2 + i\epsilon^{\alpha\mu\beta\nu} \epsilon_\alpha^{\lambda *} p_{1\mu} \epsilon_\beta^\lambda p_{2\nu} \}. \tag{7.52}$$

$\lambda = 0$

$$\left\{ -E^2 - E^2 + 2E^2 \right\} = 0. \tag{7.53}$$

$\lambda = +1$

$$\left\{ 2E^2 + i\left(\epsilon^{1023} \frac{iE^2}{2} + \epsilon^{1320} \frac{-iE^2}{2} + \epsilon^{2013} \frac{-iE^2}{2} + \epsilon^{2310} \frac{iE^2}{2} \right) \right\} = 0. \tag{7.54}$$

$\lambda = -1$

$$\left\{ 2E^2 - i\left(\epsilon^{1023} \frac{iE^2}{2} + \epsilon^{1320} \frac{-iE^2}{2} + \epsilon^{2013} \frac{-iE^2}{2} + \epsilon^{2310} \frac{iE^2}{2} \right) \right\} = 4E^2. \tag{7.55}$$

The antisymmetric four-dimensional tensor does not contribute to $\lambda = 0$, but has four independent parts for $\lambda = \pm 1$, that we have displayed separately. Reversing the sign of $\lambda = \pm 1$ changes the signs of all four terms in the expression. If in Equation (7.49) we had ϵ^{λ} for a W in the initial state rather than $\epsilon^{\lambda *}$ for a W in the final state, the polarization vector would flip. All three particles would be in the initial state, with a vacuum in the final state, so the total spin would be zero. So, in case you doubted it, the spin one polarization vectors defined by Equation (7.25) actually work. Equation (7.52) can be used to calculate a cross-section for inclusive W production:

$$\frac{1}{4}\sum_{\text{spins}} |\left\langle W^{+}\right| T \left|u\bar{d}\right\rangle|^2 = \frac{2G_F M_W^4 \cos^2\theta_c}{\sqrt{2}}. \tag{7.56}$$

In the parton model the W^+ has no transverse momentum, so it is useful to parametrize the cross-section in terms of the W rapidity:

$$y_W = \frac{1}{2}\ln\frac{E_W + P_W}{E_W - P_W}, \quad \cosh y_W = E_W/M_W, \quad \sinh y_W = P_W/M_W. \tag{7.57}$$

$M_W = 80\,\text{GeV}$, so the approximation $M_W = 0$ is not usually a good one. The initial parton momenta are related to M_W: $x_1 x_2 s = M_W^2$, and the momentum fraction of the W is $x_W = x_1 - x_2$. At 2 TeV the proton mass can be neglected, so $\sqrt{s} = 2p$. Then

$$\begin{aligned} x_W &= P_W/p = 2P_W/\sqrt{s} = \frac{2M_W}{\sqrt{s}}\sinh y_W = \frac{M_W}{\sqrt{s}}(e^{y_W} - e^{-y_W}) \\ &= x_1 - x_2. \end{aligned} \tag{7.58}$$

Hence

$$\begin{aligned} x_1 &= \frac{M_W}{\sqrt{s}}e^{y_W}, \quad x_2 = \frac{M_W}{\sqrt{s}}e^{-y_W}, \quad \pm y_{\max} = \pm\ln\left(\frac{\sqrt{s}}{M_W}\right) = \pm 3.2 \\ &\text{at 2 TeV}. \end{aligned} \tag{7.59}$$

These formulas make it easy to calculate x_1 and x_2 for a computer call to the PDF's; simply pick $-y_{\max} \le y_W \le +y_{\max}$. So, we write the parton cross-section as

$$d\hat{\sigma}(u\bar{d} \to W^+) = \frac{2G_F M_W^4 \cos^2\theta_c}{\sqrt{2}}(2\pi)^4\delta^4(p_1 + p_2 - p_3) \times \frac{d^3p_3}{(2\pi)^3 2E_3}\frac{1}{2M_W^2}. \tag{7.60}$$

There is only one particle in the final state — the W^+. The last factor is the flux, which can be evaluated in the W rest frame: $(4(E_2|\vec{p}_1| + E_1|\vec{p}_2|))^{-1} = 1/(8E^2) = 1/(2M_W^2)$. The $\int d^3p_3\delta(\vec{p}_1 + \vec{p}_2 - \vec{p}_3) = 1$, and $\delta(E_1 + E_2 - E_3) = 2E_3\delta(\hat{s} - M_W^2)$. We are left with the cross-section to be embedded in the parton model formula:

$$d\hat{\sigma}(u\bar{d} \to W^+) = \frac{2\pi G_F M_W^2 \cos^2\theta_c}{\sqrt{2}}\delta(\hat{s} - M_W^2). \tag{7.61}$$

Applying Equation (7.41) gives:

$$d\sigma = \frac{K}{3}\frac{2\pi G_F M_W^2 \cos^2\theta_c}{\sqrt{2}}\int_0^1 dx_1 \int_0^1 dx_2 (u(x_1)d(x_2) + \bar{d}(x_1)\bar{u}(x_2))\delta(\hat{s} - M_W^2). \tag{7.62}$$

To change variables, use $\hat{s} = x_1x_2s$, and $y_W = (ln(x_1) - ln(x_2))/2$. Then the Jacobian is

$$dx_1dx_2 = \frac{d\hat{s}dy_W}{s}. \tag{7.63}$$

The integral over $\hat{s}$ takes care of the δ function, and dividing by dy_W gives a differential cross-section:

$$\frac{d\sigma}{dy_W} = \frac{2\pi}{3}K\frac{G_F}{\sqrt{2}}\cos^2(\theta_c)x_1x_2(u(x_1)d(x_2) + \bar{d}(x_1)\bar{u}(x_2)), \tag{7.64}$$

where x_1 and x_2 are functions of the W^+ rapidity y_W. The constant is $15 \times 10^{-33}\,\text{cm}^2$, larger than the measured $W \to \mu\nu$ cross-section, because it does not include the leptonic branching ratio. The rapidity integral is probably a factor of three or so.

The next topic in the W boson section is the calculation of the line width Γ_W, that involves summing up all of the partial decay widths.

Table 7.1: W^+ Decay modes.

Name	Branching ratio (%)	Partial width in GeV
$l^+\nu$	(10.86 ± 0.09)	0.226
$e^+\nu$	(10.71 ± 0.16)	0.223
$\mu^+\nu$	(10.63 ± 0.15)	0.222
$\tau^+\nu$	(11.39 ± 0.21)	0.237
Hadrons	(67.41 ± 0.27)	1.405

The Particle Data Group lists the following branching ratios as in Table 7.1 for a total width $\Gamma_W = 2.085 \pm 0.042\,\text{GeV}$.

The top number should be the average of the next three numbers, and the last four numbers should add up to 100%, and both of these claims are true to a few tenths of a %. In our calculations the τ is treated as massless, just like the other leptons, while in fact $m_\tau/M_W = 0.02$. Two thirds of the W decays go into hadrons, that is the various quark-antiquark final states. Enumeration of these states requires us to extend the idea of the Cabibbo mixing angle for d and s quarks to include the b quark, giving a 3×3 matrix and substantial complication. We defer that, and work out the leptonic rates first.

To calculate the decay rates, take the back half of Equation (7.24), again dropping the BW:

$$\langle \mu^+ \nu_\mu | \, T \, | W \rangle = \frac{g}{2\sqrt{2}} \epsilon_\nu^\lambda \bar{u}(p_3) \gamma^\nu (1 - \gamma^5) v(p_4), \tag{7.65}$$

then

$$\sum_{\text{spins}} | \langle \mu^+ \nu_\mu | \, T \, | W \rangle |^2 = \frac{g^2}{8} \sum_\lambda \epsilon_\nu^\lambda \epsilon_\mu^{\lambda *} \text{Tr}(\gamma^\nu (1 - \gamma^5) \not{p}_4 \gamma^\mu (1 - \gamma^5) \not{p}_3),$$

which becomes:

$$g^2(-g_{\mu\nu} + p_\mu p_\nu / M_W^2)(g^{\alpha\mu} g^{\beta\nu} + g^{\alpha\nu} g^{\beta\mu} - g^{\alpha\beta} g^{\mu\nu} + i\epsilon^{\alpha\mu\beta\nu}) p_{3\alpha} p_{4\beta}. \tag{7.66}$$

Here $p = p_3 + p_4$ is the four momentum of the W. In the W rest frame $p = (M_W, 0, 0, 0)$. So, in the W rest frame, averaging over W

spin directions, we have:

$$\frac{1}{3}\sum_{\text{spins}} |\left\langle \mu^+\nu_\mu \right| T \left| W \right\rangle|^2 = \frac{g^2 M_W^2}{3} = \frac{8G_F M_W^4}{3\sqrt{2}}, \tag{7.67}$$

where we have used $G_F/\sqrt{2} = g^2/(8M_W^2)$. Adopting Equation (4.13) for our problem, we have the decay rate:

$$d\Gamma(\mu\nu) = \frac{8G_F M_W^4}{3\sqrt{2}} \frac{d\Omega}{64\pi^2 M_W}, \quad \Gamma(\mu\nu) = \frac{G_F M_W^3}{6\pi\sqrt{2}} = 0.22 \text{ GeV}. \tag{7.68}$$

Not bad, compared to the experimental number quoted above. In fact, we hit it right on the nose! This branching ratio should hold for $(e\nu)$, $(\mu\nu)$, and $(\tau\nu)$ decays. At the Tevatron $p\bar{p}$ collider the $(e\nu)$ and $(\mu\nu)$ modes are the principal signatures for W boson production — high p_T charged lepton plus $\not{E}_T$, so a measurement of the ratio of these two decay rates is straightforward, provided that the acceptances and efficiencies of the detector are understood. The τ lepton represents a special problem. The τ lifetime is only 0.29×10^{-12} s, $c\tau = 90\,\mu$m, about 1/5 of the B lifetime (see the discussion of b tagging in the Section 7.6). All of the τ leptons decay in a few mm, and leave no track in the detector. Since $\tau \to \nu_\tau + X$, there are two τ neutrinos in the event, and the visible energy from W decay is substantially suppressed. One interesting bias comes from the τ polarization. The τ^- is left handed, so the ν_τ tends to be in the same direction as the parent. In a decay like $\tau^- \to \pi^-\nu_\tau$ this effect gives slow pions, that are difficult to detect. Techniques have been developed to overcome these disadvantages, and the decay $W \to \tau\nu_\tau$ has been studied.

7.5. Quark Mixing in Coupling to $W^\pm$

There are two quark pairs (u, d) and (c, s), plus the b quark accessible to the 80 GeV W^+. The top quark at 175 GeV is too heavy for W decay. So, the six quark final states are $W^+ \to u\bar{d}$, $u\bar{s}$, $u\bar{b}$, $c\bar{d}$, $c\bar{s}$, and $c\bar{b}$. According to our previous assignment of the quarks into weak isospin doublets, and the favoring of transitions within the doublet, coupled with our already extensively used Cabibbo angle of $\theta_c = 13^\circ$, we anticipate that only two decay rates, into $u\bar{d}$ and $c\bar{s}$, will not be suppressed. There is a factor of three for color, so only considering these two q-qbar final states, we would guess $100\% = (3+6) \times 10.7\%$; three leptons, two quarks, and three colors. So, you see that we are only 3.7% low, and this small amount should be made up of the four forbidden decay channels.

Quark mixing is described by a 3 × 3 unitary matrix, called the CKM matrix after Cabibbo [9], and Kobayashi and Maskawa [10]. The nine entries in this matrix are fundamental parameters of the electroweak theory, and are the subject of intense experimental and theoretical study. The elements are labeled as follows:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}; \quad \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{7.69}$$

The weak interaction couples the primed charge $-1/3$ quarks to the charge $+2/3$ (u, c, t) quarks. The diagonal elements of the V_{CKM} matrix are near unity, and the bottom row, that involves the top quark, is energetically forbidden to W decay. The matrix mixes quarks, and only the top quark decay has been directly observed, so the top six entries have to be extracted from the decays of mesons and baryons that are the composite of quarks. Direct information could be obtained from accurate measurements of W branching ratios into heavy flavors, the topic of the present discussion. The Particle Data Group has published an excellent review of the various techniques used to obtain the magnitudes of these matrix elements, and global fits using all of the data are presented [11]. We reproduce their fits for the magnitudes of the matrix elements here:

$$\begin{pmatrix} 0.97446 \pm 0.0001 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359 \pm 0.0001 & 0.04214 \pm 0.00076 \\ 0.0085 \pm 0.00023 & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix}. \tag{7.70}$$

Very impressive precision! The up quark entries $|V_{ud}|$ and $|V_{us}|$ match $\cos\theta_c$ and $\sin\theta_c$. $|V_{ub}|$ is almost two orders of magnitude smaller than $|V_{us}|$. $|V_{ub}|$ would describe the decay $B^+ \to \pi^0\mu^+\nu_\mu$. $|V_{cd}|$ and $|V_{us}|$ are the same. The bottom row does not contribute to W decay. Summing the squares of the other six terms and multiplying by three for color, we are still about 3.7% low. Perhaps QCD corrections to the hadronic decays account for the difference. Direct measurements of W decays into the six light and heavy quark final states would supply useful data regarding these matrix elements, but are not easy to do, because the complication of quark fragmentation makes it difficult to tag the flavors of both quarks. An experiment to determine $|V_{cs}|$ by this method did not yield results of competitive accuracy. Heavy flavor tagging for $W^+ \to c\bar{s}$ requires reconstruction of a D meson on one side and a K meson on the other, in an event where the presence of a W^+ is assured, like in top quark decay. In any case, we have treated all of the final state particles, leptons and quarks, as massless, and the agreement between experiment and the electroweak theory is good. The reader might be surprised by the accuracy quoted for $|V_{td}|$ and $|V_{ts}|$ along the bottom row. These numbers do not come from top quark branching ratios. Only the decay $t \to W^+b$ has been observed. They come rather from so-called box diagrams, that play a role, for example, in $B\bar{B}$ mixing (see Chapter 9). The $B_s^0 \to \bar{B}_s^0$ can be modeled by an intermediate box of two W's and two t's, with vertices proportional to $|V_{ts}|$, and the parameters to describe this mixing have been measured.

CP violation in the weak interaction is described by a single complex phase in the CKM matrix. CP violation was first discovered in the $K^0\bar{K}^0$, K_SK_L system, discussed briefly in Chapter 1, but not considered in detail. There is analogous behavior in the $B\bar{B}$ system, all caused by this one complex number. CP violating effects are small, so the phase is expected also to be small, and there is some arbitrariness in its placement in the matrix. A parametrization of V_{CKM} proposed by Wolfenstein [12] is widely used, and includes the complex phase in a power series expansion of the exponential:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{7.71}$$

This matrix is unitary: $V^\dagger = V^{-1}$, if terms of order λ^4 and higher are neglected. Apart from the complex numbers, this parametrization is consistent with the magnitudes in Equation (7.70). A comparison of the two matrices gives the constant $A = 0.84$. $\lambda = 0.224$ is the Cabibbo angle in radians. The best fits for the other numbers are: $\rho = 0.14$ and $\eta = 0.36$ from the article by the Particle Data Group. CP violation results if $\eta \neq 0$. A discussion of CP violation is presented in Chapter 9.

7.6. Top Quark

The top quark is still a very active subject of research. It is so unique that it deserves a section all to itself. The top quark was discovered at the Tevatron in 1995. An early CDF top quark candidate event is shown in Figure 7.10. In $p\bar{p}$ collisions at 2 TeV the top quarks are produced in pairs by a quark antiquark annihilation and an intermediate gluon, as shown in Figure 7.11. The final state is rather complicated:

$$\begin{aligned}&p\bar{p} \to t\bar{t} + X;\ t \to W^+ b;\ \bar{t} \to W^- \bar{b};\ W^+ \to \mu^+ \nu_\mu;\ W^- \to e^- \bar{\nu}_e;\\ &b \to \bar{B};\ \bar{b} \to B.\end{aligned} \tag{7.72}$$

The cross-section for $p\bar{p} \to q\bar{q} \to g \to t\bar{t} + X$ is about 7 pb at 2 TeV, increasing rapidly to 175 pb at the LHC at 7 TeV, where the dominant production process $gg \to g \to t\bar{t}$ rather than $q\bar{q}$. The world average for the top quark mass (PDG) is 173 GeV, more than twice M_W. Hence unlike the other five quark flavors, the W boson cannot decay into a top quark: $W^+ \neq t\bar{b}$. The W cannot decay into a top, but the top can decay into a W, and the T matrix amplitude is similar to Equation (7.49):

$$-i \langle Wb| T |t\rangle = \frac{-igV_{tb}}{2\sqrt{2}} \epsilon_\mu^{\lambda *} \bar{u}(p_2)\gamma^\mu (1-\gamma^5) u(p_1), \tag{7.73}$$

where V_{tb} is the CKM matrix element in the lower right corner, ϵ_μ^λ is the W polarization vector, p_1 and p_2 are the top and b quark four momenta, respectively. The b quark mass can be neglected, but not the top quark! The evaluation of the spin averaged square, divided

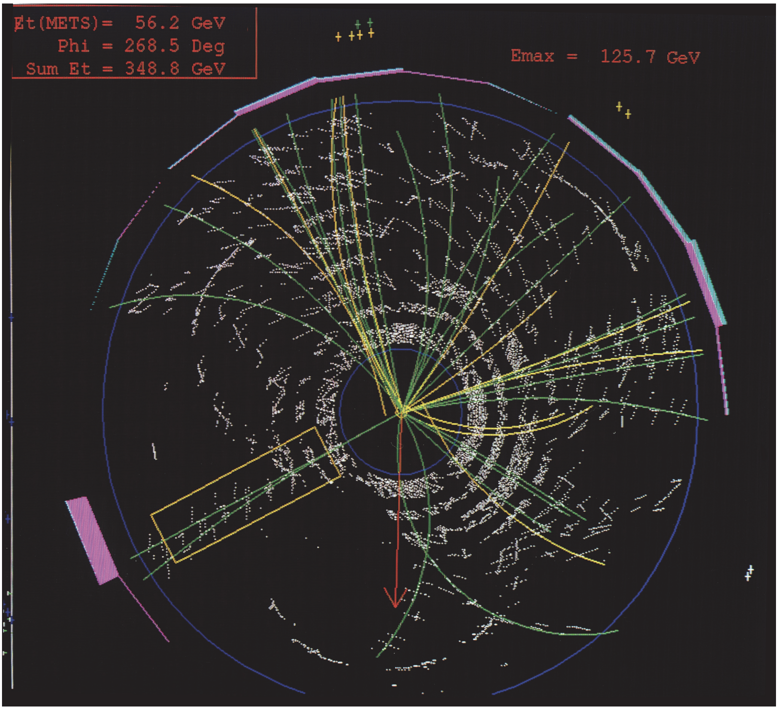


Figure 7.10: CDF event display of a top quark pair candidate. There are two oppositely charged high pt leptons in the event. One is an electron in the box below left. The red tower at about 8 o'clock is energy in the electromagnetic calorimeter. The other is a very stiff muon track heading towards 2 o'clock, about 160° clockwise from the electron. Data from the vertex detector that did the b-tagging very close to the beam pipe are not shown. The central tracking chamber has a 1-m radius, and the magnetic field is 1.4 T normal to the page. The $\not{E}_T$ vector reconstructed from the energy in the calorimeter towers, shown as color blocs at the outer edge of the tracking chamber, and the momentum of the muon, is the red arrow pointing downward. Credit: Fermilab.

by two for the top quark initial spins, is similar to W production that we have worked out before:

$$\frac{1}{2}\sum_{\text{spins}} |\langle Wb|\,T\,|t\rangle|^2 = \frac{g^2|V_{tb}|^2}{16}\sum_{\lambda} \epsilon_\nu^\lambda \epsilon_\mu^{\lambda *}\mathrm{Tr}\{(\not{p}_1 + m_1)\gamma^\nu \times (1-\gamma^5)\not{p}_2\gamma^\mu(1-\gamma^5)\}. \tag{7.74}$$

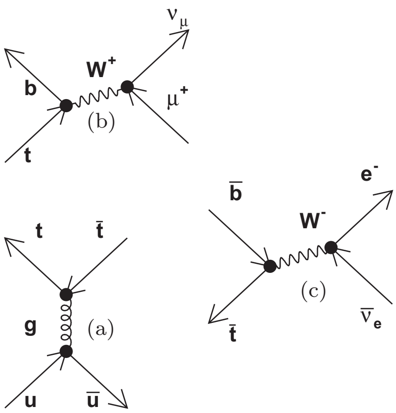


Figure 7.11: Feynman diagrams for top quark pair production and decay. Part (a) shows $u\bar{u}$ annihilation to form an intermediate gluon, but any quark–antiquark pair, or a pair of gluons, will work. The decay of top and antitop are shown separately in (b) and (c). The pattern chosen, μ^+e^-, is a $W^\pm$ signal especially free of background, but muon pairs and electron pairs are included in the sample. The neutrinos carry off missing E_T. The fragmentation of the b and $\bar{b}$ quarks into $\bar{B}$ and B mesons, respectively, is not shown, but it is the B mesons in the final state that are an important part of the experimental signature.

The completeness relation for the W polarization vector is given by Equation (7.33). We are working in the top quark rest frame: $p_1 = (m_1, 0, 0, 0)$.

$$p_1 = p_2 + p_3; \quad m_1 = m_t; \quad m_2 = m_b \sim 0; \quad m_3 = M_W;$$
$$p_1 \cdot p_2 = p_2 \cdot p_3 = m_1 E_2; \quad p_1 \cdot p_3 = m_1 E_3;$$
$$E_2 = |\vec{p}_2| = |\vec{p}_3| = \frac{m_1^2 - m_3^2}{2m_1}; \quad E_3 = \frac{m_1^2 + m_3^2}{2m_1}. \tag{7.75}$$

Neither the mass term nor the γ^5 term contributes to the trace. The result is

$$\frac{1}{2}\sum_{\text{spins}} |\langle Wb| T |t\rangle|^2 = \frac{g^2|V_{tb}|^2}{2}\left(2p_1 \cdot p_2 \right.$$
$$\left. + \frac{1}{m_3^2}(2p_1 \cdot p_3 p_2 \cdot p_3 - p_3^2 p_1 \cdot p_2)\right). \tag{7.76}$$

After using the kinematic formulas this becomes:

$$\frac{1}{2}\sum_{\text{spins}} |\langle Wb|\,T\,|t\rangle|^2 = \frac{g^2|V_{tb}|^2}{4}(m_1^2 - m_3^2)\left(2 + \frac{m_1^2}{m_3^2}\right)$$

$$= \frac{2G_F|V_{tb}|^2}{\sqrt{2}}(m_1^2 - m_3^2)(2m_3^2 + m_1^2). \quad (7.77)$$

The phase space factor for the decay rate formula is

$$d\Gamma_{i\to f} = |\langle f|\,T\,|i\rangle|^2 \frac{(2\pi)^4}{2m_1}\delta(m_1 - E_2 - E_3)\delta(\vec{p}_2 + \vec{p}_3)\frac{d^3p_2}{(2\pi)^3 2E_2} \times \frac{d^3p_3}{(2\pi)^3 2E_3}. \quad (7.78)$$

Integrating out the δ functions and integrating over the solid angle gives the total decay rate for the top quark:

$$\Gamma(t \to Wb) = \frac{G_F}{8\pi\sqrt{2}} m_t^3 \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + \frac{2M_W^2}{m_t^2}\right) = 1.92\ \text{GeV}. \quad (7.79)$$

Note that the W boson mass gets a capital letter, while the top quark, even heavier, gets a little letter because it is just a quark. The decay rate is comparable to that of the W boson, corresponding to a lifetime for the top quark of about 3×10^{-25} s. This parameter has been measured from the cross-section for single top production (see below), and from a careful study of the kinematics of $t\bar{t}$ production near threshold. A light signal does not make it across a proton in one top quark lifetime. So, the top quark does not live long enough to form mesons, or to form $t\bar{t}$ bound states — toponium. When people say that they have never seen a free quark with charge +2/3, that simply means that they have never seen a top quark, because it decays as a charge +2/3 spin 1/2 particle carrying color. There are three quark colors, and the color flows onto the daughter b quark. The b quark in turn hadronizes to make colorless B mesons that decay with lifetimes 1.5×10^{-12} s. For a B^0 meson $c\tau = 0.45$ mm, and with a Lorentz factor to help this gives a gap between two vertices of a few mm. The two vertices in the transverse plane are the beam interaction point, and the point of decay of the B meson into three or more charged tracks. This can be measured with precision tracking near the beam pipe, and furnishes a signature of heavy flavor, called a 'b tag'. There are tracks from the primary vertex, and tracks from a secondary vertex interpreted as B meson decay.

One signature for $t\bar{t}$ production is two oppositely charged high p_t leptons, electrons or muons, large $\not{E}_T$, and one or two b tags. The leptons and the $\not{E}_T$ are from the $W's$, and the b tags from the B mesons. In a real event there are two neutrinos, but there is only one $\not{E}_T$ vector, and of course it is possible for the two neutrinos to cancel each other out, leaving $\not{E}_T$ small. The two lepton, two b quark tagged, two neutrino events are the cleanest, but are not mass constrained, because of the two missing neutrinos, and the unknown energy of the produced $t\bar{t}$ pair. Mass constraints come from events in which one $W \to \mu\nu$ or $e\nu$ and the other $W \to q\bar{q}$. For these events the W mass can be reconstructed from two of the four quark jets, giving a W four momentum vector, and together with two tagged b jets one can solve for the missing neutrino, and obtain a top quark mass, assuming a $t\bar{t}$ event with $m_t = m_{\bar{t}}$. The kinematics near threshold of a large data sample of these reconstructed events is sensitive to the top quark width, and one measurement of this parameter comes from such an analysis [13]. Single top production can occur through an intermediate off-shell $W \to t\bar{b}$, an s channel Drell–Yan process, or through t channel W exchange $q + b \to q' + t$, where the b quark comes from the proton sea. The Tevatron cross-section $\sigma(p\bar{p} \to t + X)$ is about 3 pb, half as large as the pair cross-section, with the ratio (s channel)/(t channel) = 1/2. Single top grows to about 64 pb at 7 TeV at the LHC. This cross-section is of interest because it is proportional to the top quark width, Equation (7.79), usually interpreted as a measurement of the CKM matrix element $|V_{tb}|$ with the kinematic factors in Equation (7.79) removed [14]. A single top s channel event has a W boson and two b quarks, while the t channel has one b quark and one light quark. The other b quark from the proton sea is part of the underlying event. The cross-section is 0.1% of the W cross-section, so it is difficult to separate from a background of W + jets. Even the top pair cross-section represents a background. The measurements of $|V_{tb}|$ are consistent with one. There is no experimental observation of the top quark coupling to any other charge $-1/3$ quark.

The calculation of W polarization in top quark decay is similar to $u\bar{d} \to W^+$ calculated above. In fact, the calculation is slightly simpler, if you are not fond of permuting the indices of the four-dimensional antisymmetric tensor, because there are only two terms. The answer is quite different, however. Because of the large value of the top quark mass, the helicity operators that couple only left-handed quarks are not in force, and the $\lambda = 0$ component of the W polarization, the one that is absent for the photon, dominates in top quark decay. This is of interest in the electroweak theory, as we shall learn in Chapter 8 that this component of the W spin is absent from the basic Lagrangian where all particles are massless, and becomes non-zero as a result of the spontaneous symmetry breaking that generates the masses of the W and Z bosons. We begin with Equation (7.74), and write down the squared matrix element, spin averaged:

$$\begin{aligned}\frac{1}{2}\sum_{\text{spins}} |\, \langle Wb|\, T\, |t\rangle \,|^2 &= \frac{g^2|V_{tb}|^2}{16}\sum_{\lambda}\mathrm{Tr}\{\not{p}_2\not{\epsilon}^{\lambda *}(1-\gamma^5)\not{p}_1\not{\epsilon}^{\lambda}(1-\gamma^5)\}\\ &= \frac{g^2|V_{tb}|^2}{2}\sum_{\lambda}\{\epsilon^{\lambda *}\cdot p_2\epsilon^{\lambda}\cdot p_1 + \epsilon^{\lambda *}\cdot p_1\epsilon^{\lambda}\cdot p_2\\ &\quad + p_1\cdot p_2 - i\epsilon^{\mu\alpha\nu\beta}p_{2\mu}\epsilon^{\lambda *}_{\alpha}p_{1\nu}\epsilon^{\lambda}_{\beta}\}. \end{aligned} \tag{7.80}$$

We evaluate this expression in the top rest frame, with the $\hat{z}$ axis pointing in the direction of the final state W^+ boson. In this frame the lower index four vectors are

$$\begin{aligned}&p_1 = (m_t, 0, 0, 0), \quad p_2 = (E_b, 0, 0, E_b), \quad p_W = (E_W, 0, 0, -E_b),\\ &E_b = |\vec{p}| = \frac{m_t^2 - M_W^2}{2m_t}, \quad E_W = \frac{m_t^2 + M_W^2}{2m_t}, \quad p_1\cdot p_2 = m_t E_b,\\ &\epsilon^0 = (E_b, 0, 0, -E_W)/M_W = \epsilon^{0*}, \quad \epsilon^{+1} = (0, 1, i, 0)/\sqrt{2},\\ &\epsilon^{+1*} = (0, 1, -i, 0)/\sqrt{2}, \quad \epsilon^{-1} = (0, -1, i, 0)/\sqrt{2},\\ &\epsilon^{-1*} = (0, -1, -i, 0)/\sqrt{2}, \quad \epsilon^{\pm 1}\cdot p_1 = \epsilon^{\pm 1}\cdot p_2 = 0, \quad \epsilon^0\cdot p_1 = \frac{m_t E_b}{M_W},\\ &\epsilon^0\cdot p_2 = \frac{E_b^2 + E_b E_W}{M_W}. \end{aligned} \tag{7.81}$$

Assembling this array of formulas into Equation (7.80) gives apart from the overall constant:

$$\begin{aligned}
&\lambda = 0 : \left\{ m_t E_b \left(1 + \frac{2m_t E_b}{M_W^2} \right) \right\} ; \\
&\lambda = -1 : \left\{ m_t E_b \left(1 - \frac{1}{2}\epsilon^{3102} + \frac{1}{2}\epsilon^{3201} \right) \right\} = 2m_t E_b ; \\
&\lambda = +1 : \{\} = 0.
\end{aligned} \tag{7.82}$$

The signs of the antisymmetric tensor terms flip going from $\lambda = -1$ to $\lambda = +1$. Convince yourself that $\lambda = +1$ is forbidden because the b quark is left handed. The fraction of W^+ from top decay that carry $\lambda = 0$ polarization is

$$\frac{(\lambda = 0)}{(\lambda = 0) + (\lambda = -1)} = \frac{1}{1 + 2M_W^2/m_t^2} = 0.7. \tag{7.83}$$

This is a remarkable result: 70% of all W's from top decay have no helicity; the spin magnetic quantum number along the direction of the W motion is $m_S = 0$. Since the angle of emission of the charged lepton in $W^+ \to \mu^+\nu_\mu$ is correlated with the W^+ spin direction, this parameter can be measured from a sample of top decays. The measured lepton asymmetries are consistent within experimental uncertainties of these predictions [15].

Top quark studies are ongoing, and there is still much to be learned about this curious particle. We have relied here on Tevatron results, where the quark was first seen, but the LHC has much larger top production cross-sections, and new results are coming out all the time.

7.7. Neutral Currents and Neutrinos

The coupling of the neutral Z intermediate boson to fermions is more intricate than the charged current coupling to the $W^\pm$. We have already introduced the concept of weak isospin, and left-handed doublets for the fermions. The next step would be to introduce a weak isospin triplet of W bosons, with W_μ^0 the neutral candidate:

$$W_\mu^\pm = \frac{W_\mu^{(1)} \mp iW_\mu^{(2)}}{\sqrt{2}}, \quad W_\mu^0 = W_\mu^{(3)}. \tag{7.84}$$

The only problem with this assignment is that it is too restrictive, because it would require the neutral coupling W_μ^0 to be the same as the charged $V - A$ coupling, and it turns out that nature prefers a different scheme for the neutral current. To allow for this variation, an extra neutral vector field that is a weak isospin singlet, and hence couples to right-handed leptons, is introduced, called B_μ, and a mixing angle θ_W to create the Z and A (photon) fields from mixtures:

$$A_\mu = B_\mu \cos\theta_W + W_\mu^{(3)} \sin\theta_W, \quad Z_\mu = -B_\mu \sin\theta_W + W_\mu^{(3)} \cos\theta_W. \tag{7.85}$$

All four fields W_μ and B_μ start out massless, and hence satisfy gauge invariance, and after the smoke clears on the dynamical symmetry breaking (that creates the Higgs boson), the $W_\mu^\pm$ and Z_μ acquire masses, but A_μ remains massless. We will go through this argument in detail Chapter 8. For now, let us see how it works. The angle θ_W is a parameter of the theory, to be determined by experiment. It is often referred to as the Weinberg angle, in recognition of Steve Weinberg's significant contributions to the subject, but it can also be called the weak angle. 'W' works either way. We follow the conventions of the Particle Data Group in our definitions. The coupling of the Z boson to fermions is written as

$$-\frac{g}{2\cos\theta_W} \sum_i \bar{\psi}_i \gamma^\mu (g_V^i - g_A^i \gamma^5) \psi_i Z_\mu. \tag{7.86}$$

The constant g in front is the same as we have been using for the W charged current coupling, with $G_F/\sqrt{2} = g^2/(8M_W^2)$. It is related to the elementary charge by Equation (6.5): $g \sin\theta_W = e$. The state vector ψ does not have an 'L' subscript, because both left and right-handed states are allowed. The three most accurate parameters of the electroweak theory are M_Z, G_F, and α, where according to the Particle Data Group the best values are

$$M_Z = 91.1876 \pm 0.0021\,\text{GeV}; \quad G_F = 1.1663787(6) \times 10^{-5}\,\text{GeV}^{-2};$$
$$1/\alpha = 137.035999150(33). \tag{7.87}$$

The fine structure constant is known to 12 decimal places, but we will usually round it off. The precision of the other two constants is also impressive. M_Z comes from LEP, where the energy of the e^-e^+ storage ring running at the Z peak had to be very precisely known, and G_F comes from the muon lifetime. How LEP calibrated its magnetic field is a chapter in itself, but the basic idea was to use resonances based on the electron g factor. Synchrotron radiation polarizes the electron beam perpendicular to the orbit, and the beam can be depolarized by an oscillating horizontal magnetic field. The depolarization frequency is proportional to the magnetic moment anomaly and the energy of the beam. Beam polarization was measured by a Compton polarimeter. The ring energy ($M_Z/2$) at collision points was known to better than 1 MeV, consistent with the ±2 MeV quoted error for M_Z [16]. One of the many relations for the Weinberg angle is

$$\sin^2\theta_W = s_W = 1 - \frac{M_W^2}{M_Z^2} = 0.22337 \pm 0.0001. \tag{7.88}$$

This is a non-trivial relation. Since θ_W is a real angle, it requires $M_W \leq M_Z$, as confirmed by experiment. The precise value for θ_W depends upon how it is measured, an effect that is understood by radiative corrections to the first-order electroweak theory that is the subject of this text. A widely quoted value is $\sin^2\theta_W = 0.231$ that comes from asymmetry measurements with Z's produced by polarized electrons [17]. However, we will use the PDG value in calculations here, and also often use the notation $s_W = \sin^2\theta_W$. The Z couplings to fermions in the Lagrangian Equation (7.86) are as follows:

$$\begin{aligned}
&g_V^i = t_{3L}(i) - 2Q_i\sin^2\theta_W; \quad g_A^i = t_{3L}(i);\\
&t_{3L}(i) = +1/2, \nu_e, \nu_\mu, \nu_\tau, u, c, t,\\
&t_{3L}(i) = -1/2, e^-, \mu^-, \tau^-, d, s, b;\\
&Q_i = 0, \nu'\text{s}; \quad Q_i = -1, e^-, \mu^-, \tau^-,\\
&Q_i = +2/3, u, c, t; \quad Q_i = -1/3, d, s, b.
\end{aligned} \tag{7.89}$$

These equations have several interesting features. First of all, if $Q_i = 0$, the neutrinos, then the couplings to the $W_\mu^\pm$ and Z_μ are the

Table 7.2: Electroweak coupling of the Z to fermions.

Fermion	t_{3L}	Q	$g_V(s_W)$	g_V	g_A	g_L	g_R	A^f
ν_e, ν_μ, ν_τ	$+1/2$	0	$+1/2$	$+1/2$	$+1/2$	$1/2$	0	1
e, μ, τ	$-1/2$	-1	$-1/2 + 2s_W$	-0.054	$-1/2$	-0.277	0.223	0.215
u, c, t	$+1/2$	$+2/3$	$+1/2 - (4/3)s_W$	0.203	$+1/2$	0.351	0.148	0.70
d, s, b	$-1/2$	$-1/3$	$-1/2 + (2/3)s_W$	-0.351	$-1/2$	-0.425	0.074	0.94

same, and this is almost true for the charge $-1/3$ quarks. If $Q_i = -1$, and $\sin^2 \theta_W = 0.25$, or $\theta_W = 30°$, which is also almost true, then the vector coupling vanishes. The coupling of the Z_μ to e^-, μ^-, τ^- is almost parity conserving pure axial vector, and of course the coupling to A_μ is pure vector. We will often find it useful to work with the orthogonal helicity states rather than vector and axial vector. They are defined as follows:

$$(g_V - g_A\gamma^5)\gamma^\mu = g_L(1-\gamma^5)\gamma^\mu + g_R(1+\gamma^5)\gamma^\mu; \quad g_V = g_L + g_R;$$
$$g_A = g_L - g_R. \tag{7.90}$$

Because $(1 \pm \gamma^5)(1 \pm \gamma^5) = 2(1 \pm \gamma^5)$, and $(1 \pm \gamma^5)(1 \mp \gamma^5) = 0$, this choice often simplifies the trace calculations.

These numbers can be summarized in Table 7.2.

The fifth column g_V has $s_W = 0.223$ substituted into the fourth column g_V formula. Columns seven and eight are the left-handed and right-handed couplings, $g_L = (g_V + g_A)/2$, and $g_R = (g_V - g_A)/2$, and the last column is the asymmetry parameter $A^f = ((g_L^f)^2 - (g_R^f)^2)/((g_L^f)^2 + (g_R^f)^2)$. The asymmetry parameters play a key role in the discussion of Z production by polarized electrons below.

We have a new interaction that depends on the charge and third component of weak isospin of the fermions. Perhaps the simplest problem to solve is $\nu_\mu e^- \to \nu_\mu e^-$, a process that can only occur via the weak neutral current. There is also $\nu + \nu \to \nu + \nu$, that may have astrophysical interest with the cosmic neutrino background, but is otherwise difficult to measure. The reaction $\nu_e + e^- \to \nu_e + e^-$ has already been considered as an example of t channel W exchange, a diagram that is forbidden for a muon neutrino. The neutral current can also contribute to $\nu_e + e^- \to \nu_e + e^-$, see Problem 7.6.

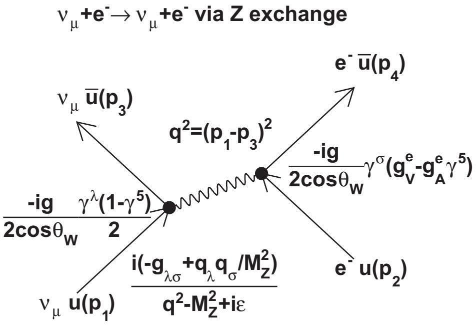

$\nu_\mu + e^- \to \nu_\mu + e^-$ via Z exchange
$e^-\ \bar{u}(p_4)$
$\nu_\mu\ \bar{u}(p_3)$
$q^2=(p_1-p_3)^2$
$\frac{-ig}{2\cos\theta_W}\gamma^\sigma(g_V^e-g_A^e\gamma^5)$
$\frac{-ig}{2\cos\theta_W}\frac{\gamma^\lambda(1-\gamma^5)}{2}$
$\frac{i(-g_{\lambda\sigma}+q_\lambda q_\sigma/M_Z^2)}{q^2-M_Z^2+i\varepsilon}$
$e^-\ u(p_2)$
$\nu_\mu\ u(p_1)$

Figure 7.12: Scattering of ν_μ by electrons via neutral current t channel exchange. The Z couplings to spin 1/2 Dirac fermions are listed in Table 7.2. θ_W is the Weinberg angle.

But first, back to the scattering of muon neutrinos by electrons, shown in Figure 7.12. This is our first example of a diagram where the interactions at the two vertices are different. The Z propagator has the same form as the W propagator in Equation (7.24):

$$\begin{aligned}-i\left\langle \nu_\mu e^- \right| T \left| \nu_\mu e^- \right\rangle &= \left(\frac{-g}{2\cos\theta_W}\right)^2 \bar{u}(p_3)\gamma^\mu \frac{(1-\gamma^5)}{2} u(p_1)\\ &\quad\times\left\{\frac{i(-g_{\mu\nu}+q_\mu q_\nu/M_Z^2)}{q^2-M_Z^2+iM_Z\Gamma_z}\right\}\\ &\quad\times \bar{u}(p_4)\gamma^\nu(g_V-g_A\gamma^5)u(p_2);\\ p_1+p_2 &= p_3+p_4;\ q=p_1-p_3. \end{aligned} \tag{7.91}$$

The Mandelstam variables are $s=(p_1+p_2)^2$, $t=(p_1-p_3)^2$, and $u=(p_1-p_4)^2$. Table 7.2 gives $g_V=-0.054$ and $g_A=-1/2$, but we will just leave the variables in the equation until the end. For $|q^2| \ll M_Z^2$ we may replace the propagator with $ig_{\mu\nu}/M_Z^2$, leaving the expression:

$$\begin{aligned}\left\langle \nu_\mu e^- \right| T \left| \nu_\mu e^- \right\rangle &= \frac{g^2}{8\cos^2\theta_W M_Z^2}\bar{u}(p_3)\gamma^\mu(1-\gamma^5)u(p_1)\bar{u}(p_4)\gamma_\mu\\ &\quad\times(g_V-g_A\gamma^5)u(p_2).\end{aligned} \tag{7.92}$$

It is convenient to write the electron term as a function of g_L and g_R, thus,

$$\begin{aligned}\left\langle \nu_\mu e^- \right| T \left|\nu_\mu e^-\right\rangle &= \frac{g^2}{8\cos^2\theta_W M_Z^2}\bar{u}(p_3)\gamma^\mu(1-\gamma^5)u(p_1)\bar{u}(p_4)\gamma_\mu \\ &\quad \times(g_L(1-\gamma^5)+g_R(1+\gamma^5)u(p_2) \\ &= \frac{g^2}{8\cos^2\theta_W M_Z^2}(g_L(\bar{u}(p_3)\gamma^\mu(1-\gamma^5) \\ &\quad \times u(p_1)\bar{u}(p_4)\gamma_\mu(1-\gamma^5)u(p_2)) \\ &\quad + g_R(\bar{u}(p_3)\gamma^\mu(1-\gamma^5)u(p_1)\bar{u}(p_4)\gamma_\mu(1+\gamma^5)u(p_2)) \\ &= \frac{g^2}{8\cos^2\theta_W M_Z^2}(g_L A + g_R B). \end{aligned} \tag{7.93}$$

The first term, A, proportional to g_L, looks familiar, the second term, B, proportional to g_R, is not the same, but not very different either. The trace for the $(V-A)\times(V-A)$ term has been evaluated in Equation (6.36), Section 6.2 of Chapter 6 on muon decay, and you may convince yourself that the cross terms like $AB^\dagger$ vanish, because $(1-\gamma^5)(1+\gamma^5)=0$. Therefore, we have

$$\begin{aligned}\frac{g_L^2}{2}\sum_{\text{spins}} AA^\dagger &= 2g_L^2\mathrm{Tr}(\not{p}_3\gamma^\mu(1-\gamma^5)\not{p}_1\gamma^\sigma)\mathrm{Tr}(\not{p}_4\gamma_\mu(1-\gamma^5)\not{p}_2\gamma_\sigma) \\ &= 128 g_L^2(p_1\cdot p_2)(p_3\cdot p_4), \end{aligned}\tag{7.94}$$

and

$$\begin{aligned}\frac{g_R^2}{2}\sum_{\text{spins}} BB^\dagger &= 2g_R^2\mathrm{Tr}(\not{p}_3\gamma^\mu(1-\gamma^5)\not{p}_1\gamma^\sigma)\mathrm{Tr}(\not{p}_4\gamma_\mu(1+\gamma^5)\not{p}_2\gamma_\sigma) \\ &= 128 g_R^2(p_1\cdot p_4)(p_2\cdot p_3). \end{aligned}\tag{7.95}$$

Evaluation of $BB^\dagger$ and $AA^\dagger$ are the same, only a sign changes for the $\epsilon^{\mu\alpha\nu\beta}\epsilon_{\mu\rho\nu\sigma}$ term in Equation (6.36). $AA^\dagger$ is $(V-A)\times(V-A)$, and Barger and Phillips say that both $(V-A)\times(V-A)$ and $(V+A)(V+A)$ have incoming arrows dotted together, and outgoing arrows dotted together. $BB^\dagger$ is $(V-A)\times(V+A)$, and for this case the incoming arrow on one side of the exchanged Z is dotted into the outgoing arrow on the other side. A glance at the diagram in Figure 7.12 shows that this is indeed the case. Final assembly gives for the squared averaged matrix element:

$$\frac{1}{2}\sum_{\text{spins}} |\left\langle \nu_\mu e^- \right| T \left| \nu_\mu e^- \right\rangle|^2 = \frac{g^4}{2\cos^4\theta_W M_Z^4}(g_L^2 s^2 + g_R^2 u^2). \tag{7.96}$$

Here, we have neglected the electron mass. For the charged current cross-section $\sigma(\nu_e e^- \to \nu_e e^-)$ Equation (7.5), which is pure $(V - A)$, the mass does not contribute. But here, for the neutral current, the mass does contribute if there are right-handed terms. Retaining the electron mass gives an extra term $2m^2 g_L g_R t$ to be added in the parentheses of Equation (7.96). This term is small, even for solar neutrinos. See Problems 7.2–7.4. Without the mass term this formula is usually written in terms of the variable y introduced in Equation (5.84). For our neutrino scattering, the definition is

$$y = \frac{2p_2 \cdot (p_1 - p_3)}{2p_1 \cdot p_2} = 1 - \frac{p_2 \cdot p_3}{p_1 \cdot p_2} = 1 + \frac{u}{s},\ 0 \le y \le 1. \tag{7.97}$$

So,

$$\begin{aligned}\frac{1}{2}\sum_{\text{spins}} |\left\langle \nu_\mu e^- \right| T \left| \nu_\mu e^- \right\rangle|^2 &= \frac{g^4 s^2}{2\cos^4\theta_W M_Z^4}(g_L^2 + g_R^2(1-y)^2) \\ &= 16 G_F^2 s^2 (g_L^2 + g_R^2(1-y)^2),\end{aligned} \tag{7.98}$$

where we have used $G_F/\sqrt{2} = g^2/(8M_W^2)$, and $\cos\theta_W = M_W/M_Z$. Now use $d\Omega = 4\pi dy$ and Equation (4.69) for the cross-section:

$$\begin{aligned}\frac{d\sigma(\nu_\mu e^- \to \nu_\mu e^-)}{dy} &= \frac{G_F^2 s}{\pi}(g_L^2 + g_R^2(1-y)^2) = \frac{G_F^2 s}{4\pi}((g_V + g_A)^2 \\ &\quad + (g_V - g_A)^2(1-y)^2).\end{aligned} \tag{7.99}$$

If the electron were replaced by an electron neutrino, difficult to do experimentally, but this is a thought experiment, then $g_V = g_A = 1/2$, the $(1 - y)^2$ term vanishes, and we are left with a total cross-section $\sigma = G_F^2 s/(4\pi)$, which looks just like a charged current neutrino cross-section. The $(1 - y)^2$ term comes from right-handed electrons, that are allowed by the neutral current in the electroweak theory. The angle in $u = (p_1 - p_4)^2$ is between the incident left-handed neutrino and the outgoing electron, which cannot be right-handed in the forward direction. The total cross-section is

$$\sigma(\nu_\mu e^- \to \nu_\mu e^-) = \frac{G_F^2 s}{\pi}(g_L^2 + g_R^2/3) = 1.5 \times 10^{-42}\text{cm}^2 \times E_\nu \text{ in GeV.} \tag{7.100}$$

This number is about 10% of Equation (7.5) for the charged current cross-section $\sigma(\nu_e e^- \to \nu_e e^-) = 1.7 \times 10^{-41}\text{cm}^2 \times E_\nu$ in GeV.

The cross-section formula Equation (7.99) can be adopted for neutral current neutrino scattering in the parton model. For u and d quarks, all we need to do is substitute $s \to \hat{s} = xs$, and read the correct couplings from Table 7.2. Thus,

$$\frac{d\sigma(\nu_\mu u \to \nu_\mu u)}{dy} = \frac{G_F^2 xs}{\pi}((g_L^u)^2 + (g_R^u)^2(1-y)^2), \tag{7.101}$$

and

$$\frac{d\sigma(\nu_\mu d \to \nu_\mu d)}{dy} = \frac{G_F^2 xs}{\pi}((g_L^d)^2 + (g_R^d)^2(1-y)^2). \tag{7.102}$$

Here the muon neutrino is not special. Any flavor neutrino will have the same formula. The kinematic variables are shuffled to describe $\nu + \bar{u}$ scattering. For scattering by antiquarks, the momenta p_4 and $-p_2$ are interchanged, which interchanges $\hat{s}$ and $\hat{u}$, leaving $\hat{t}$ the same. This shifts the $(1-y)^2$ from g_R^2 to g_L^2. Thus,

$$\frac{d\sigma(\nu_\mu \bar{u} \to \nu_\mu \bar{u})}{dy} = \frac{G_F^2 xs}{\pi}((g_R^u)^2 + (g_L^u)^2(1-y)^2), \tag{7.103}$$

and

$$\frac{d\sigma(\nu_\mu \bar{d} \to \nu_\mu \bar{d})}{dy} = \frac{G_F^2 xs}{\pi}((g_R^d)^2 + (g_L^d)^2(1-y)^2). \tag{7.104}$$

The overall sign of the neutral couplings defined in Equation (7.89) changes going from particle to antiparticle, so that the squares that appear in these formulas are the same for u quarks and $\bar{u}$ quarks. For antineutrinos the relations are: $\sigma(\bar{\nu}u) = \sigma(\nu\bar{u})$, and $\sigma(\bar{\nu}\bar{u}) = \sigma(\nu u)$.

These cross-sections are then embedded in parton model formulas like Equation (7.23). For example,

$$
\begin{aligned}
\frac{d\sigma(\nu_\mu + p \to \nu_\mu + X)}{dxdy} = \frac{G_F^2 s}{\pi}\{&((g_L^u)^2 + (g_R^u)^2(1-y)^2)xu(x) \\
&+ ((g_L^d)^2 + (g_R^d)^2(1-y)^2)xd(x) \\
&+ ((g_L^u)^2(1-y)^2 + (g_R^u)^2)x\bar{u}(x) \\
&+ ((g_L^d)^2(1-y)^2 + (g_R^d)^2)x\bar{d}(x)\}. \qquad (7.105)
\end{aligned}
$$

The different couplings make the neutral current formula look more complicated than the corresponding charged current formula Equation (7.23). The small value for $(g_L^u)^2$ accounts for the claim that neutral current effects in neutrino scattering from materials are about 20% of the charged current.

7.8. *Z* Production and Decay

Figure 7.13 shows two Feynman diagrams for the s-channel process $e^-e^+ \to \mu^-\mu^+$, one with an intermediate photon, and the other with an intermediate Z boson. We do not consider the analog of Bhabha scattering $e^-e^+ \to Z \to e^-e^+$, because of the interference of

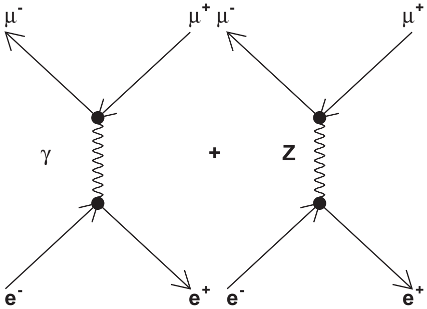


Figure 7.13: Feynman diagrams for $e^- + e^+ \to \mu^- + \mu^+$ by $\gamma - Z$ interference. The diagrams look the same, but the couplings at the two vertices are different. The γ exchange amplitude is real, and the interference with the Z Breit–Wigner changes sign through the resonance, and vanishes at $s = M_Z^2$. On resonance the square of the right-hand diagram is over 100 times larger than the left.

a t-channel exchange diagram. So, we limit our discussion to $e^-e^+ \to Z \to f\bar{f}$, where the final state fermion is not an electron. Using the Feynman rules, the two matrix elements are

$$-i\left\langle\mu^-\mu^+\right|\gamma\left|e^-e^+\right\rangle = \bar{v}(p_2)(-ie\gamma^\mu u(p_1)\left\{\frac{-ig_{\mu\nu}}{s}\right\}\bar{u}(p_3)(-ie\gamma^\nu v(p_4), \tag{7.106}$$

and

$$\begin{aligned}-i\left\langle\mu^-\mu^+\right|Z\left|e^-e^+\right\rangle = \bar{v}(p_2)&\left(\frac{-ig}{2\cos\theta_W}\gamma^\lambda(g_V^e - g_A^e\gamma^5)\right)u(p_1)\\ &\times\left\{\frac{-i(g_{\lambda\sigma} - q_\lambda q_\sigma/M_Z^2)}{s - M_Z^2 + iM_Z\Gamma_Z}\right\}\bar{u}(p_3)\\ &\times\left(\frac{-ig}{2\cos\theta_W}\gamma^\sigma(g_V^\mu - g_A^\mu\gamma^5)\right)v(p_4),\end{aligned} \tag{7.107}$$

where $q = p_1 + p_2 = p_3 + p_4$, and $s = q^2$. The mass term in the numerator of the Z propagator does not contribute because of the Dirac equation for massless spin 1/2 particles. Write it as $q_\lambda q_\sigma = (p_1 + p_2)_\lambda(p_3 + p_4)_\sigma$, then the left-hand part gives Dirac equations for the left spinors, and the right-hand part for the right spinors. So, we may neglect that term in our discussion. Adding the two amplitudes then gives:

$$\begin{aligned}\left\langle\mu^-\mu^+\right|T\left|e^-e^+\right\rangle = &-\frac{e^2}{s}\bar{v}(p_2)\gamma^\mu u(p_1)\bar{u}(p_3)\gamma_\mu v(p_4)\\ &-\frac{g^2}{4\cos^2\theta_W(s - M_Z^2 + iM_Z\Gamma_Z)}\\ &\times\bar{v}(p_2)\gamma^\lambda(g_V - g_A\gamma^5)u(p_1)\\ &\times\bar{u}(p_3)\gamma_\lambda(g_V - g_A\gamma^5)v(p_4).\end{aligned} \tag{7.108}$$

The e and μ couplings are the same, so we have dropped the superscript, and that makes the formulas less cumbersome. However, there is a price to pay, because we will have to backtrack and put the labels back in for Z production by q-qbar annihilation (Drell–Yan in $p\bar{p}$ collisions), or if we were interested in studying $Z \to q\bar{q}$ decay. In the

Drell–Yan case at $\sqrt{s} = 2$ TeV the cross-section times branching ratio for $p\bar{p} \to W^{\pm} \to \mu^{\pm}$ is about ten times as large as $p\bar{p} \to Z \to \mu^{-}\mu^{+}$. The suppression comes from two factors of three, one for the production cross-section, and the other for the branching fraction to lepton pairs. As we shall soon see, $Z \to \mu^{-}\mu^{+}$ is only about 3% of all Z decays.

Rationalize the BW:

$$\frac{1}{s - M_Z^2 + iM_Z\Gamma_Z} = \frac{s - M_Z^2 - iM_Z\Gamma_Z}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2}. \tag{7.109}$$

The electromagnetic term is real, so it is only the real part of the BW that contributes to the interference, and that term $(s - M_Z^2)$ vanishes on the resonance, and changes sign as you go through the resonance-adding on one side, and subtracting on the other. Because the linear BW falls off more slowly than the quadratic one, this interference term can be detected well below 91 GeV, giving evidence for the presence of the Z resonance before its discovery. The squared matrix element has four terms:

$$|\langle\mu\mu|\,T\,|ee\rangle|^2 = AA^{\dagger} + AB^{\dagger} + BA^{\dagger} + BB^{\dagger}. \tag{7.110}$$

The first term is the familiar photon term with cross-section one unit of R: $\sigma = (4\pi\alpha^2)/(3s)$. The next two terms are interference terms that vanish on the Z peak, but are useful below the resonance as evidence that the Z exists. The $BB^{\dagger}$ term is the one of interest — the Z resonance itself. The first term spin averaged gives Equation (4.56):

$$\frac{1}{4}\sum_{\text{spins}} AA^{\dagger} = e^4(1 + \cos^2\theta). \tag{7.111}$$

The polar angle θ is between the incident electron and the outgoing μ^{-} in the center of mass, where $\vec{p}_1 \cdot \vec{p}_3 = s\cos\theta/4$. The interference term is

$$\begin{aligned}\frac{1}{2}\sum_{\text{spins}} BA^{\dagger} = {}& \frac{8\pi^2\alpha^2}{\cos^2\theta_W \sin^2\theta_W}\frac{s(s - M_Z^2)}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \\ & \times \{(g_L + g_R)^2(1 + \cos^2\theta) + 2(g_L - g_R)^2\cos\theta\}.\end{aligned} \tag{7.112}$$

Note that the interference term has a small asymmetry proportional to $(g_L - g_R)^2$, that comes from parity violation in Z production and decay. This means that below the resonance, $\sqrt{s} = 30\,\text{GeV}$ for example, there is a front-back asymmetry in the μ^- angular distribution that comes from $\gamma - Z$ interference. The Z term is

$$\frac{1}{4}\sum_{\text{spins}} BB^\dagger = \frac{4\pi^2\alpha^2}{\cos^4\theta_W \sin^4\theta_W} \frac{s^2}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \times \{(g_L^2 + g_R^2)^2(1 + \cos^2\theta) + 2(g_L^2 - g_R^2)^2 \cos\theta\}. \tag{7.113}$$

The angular dependences of the interference term and the Z resonance term are the same, with different coefficients. Integration over the solid angle kills the $\cos\theta$ term and gives a factor $(16\pi)/3$, so the total cross-section is

$$\begin{aligned}\sigma(e^-e^+ \to \mu^-\mu^+) = \frac{4\pi\alpha^2}{3s} &+ \frac{2\pi\alpha^2}{3\cos^2\theta_W \sin^2\theta_W} \\ &\times \frac{s - M_Z^2}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2}(g_L + g_R)^2 \\ &+ \frac{\pi\alpha^2}{3\cos^4\theta_W \sin^4\theta_W} \frac{s}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \\ &\times (g_L^2 + g_R^2)^2. \end{aligned} \tag{7.114}$$

The electroweak Z coupling constant has been replaced by using $g\sin\theta_W = e$, so that each term is proportional to α^2. The Z coupling has the Weinberg angle in the denominator, making it actually larger than just α^2. So, the weak interaction is not all that weak in these formulas involving the Z propagator. On the Z peak, $s = M_Z^2$, the interference term vanishes, and the Z term is about 100 times larger than $4\pi\alpha^2/(3s)$, which is by definition one unit of R.

$$\frac{\sigma(ee \to Z \to \mu\mu)}{R} = \frac{M_Z^2}{4\Gamma_Z^2} \frac{(g_L^2 + g_R^2)^2}{\cos^4\theta_W \sin^4\theta_W} = 165. \tag{7.115}$$

Near the resonance you may safely ignore the background. Incidentally, one unit of R at the Z peak is $4\pi\alpha^2/(3M_Z^2) = 10^{-35}\,\text{cm}^2 =$ 10 pb, which is a small cross-section. So, the extra factor of 165 on the Z peak helps a whole lot. LEP, the electron–positron collider at CERN that ran on the Z peak for years, quotes an effective cross-section of 30 nb, for $Z \to$ hadrons. This is consistent with our numbers, because the branching ratio $\Gamma(Z \to q\bar{q})/\Gamma(Z \to \mu^-\mu^+) = 20$. Typical luminosity at LEP for running at $50\,\text{GeV} \times 50\,\text{GeV}$ was $2 - 3 \times 10^{31}\text{cm}^{-2}\,\text{s}^{-1}$, giving about 0.6 $Z \to$ hadrons per second, a very good event rate. If the ring is tuned to the Z resonance, every event is interesting!

There were two $e^-e^+ \to Z$ factories in operation at approximately the same time (mid to late 1990s). LEP was a storage ring, where electrons and positrons coasted around a circle 27 km in circumference. Most of the running time was devoted to 45 GeV ring energy, for the Z resonance, but some running at 90 GeV was done at the end to make W pairs. The other accelerator was the SLAC Linear Collider, or SLC, which was a remarkable device, and its detector, called SLD. There is a very good review of the SLC/SLD physics program. See [18]. The linear collider was made from the SLAC linac by accelerating bunches of electrons and positrons down the same line, bending them in opposite directions and then back for a head on collision. Nominal energy was again 45 GeV. SLC was a single pass collider, with bunches of charge passing through each other at 120 Hz. In the final running period from 1997 to 1998 the luminosity reached a peak of $\mathcal{L} = 3 \times 10^{30}\text{cm}^{-2}\text{s}^{-1}$, about a factor of 10 less than LEP. The SLC had a secret weapon, however. It could accelerate polarized electrons down the linac, and preserve the polarization throughout all of the elaborate choreography that the electrons were subject to. The polarization came from a GaAs photocathode that was irradiated by circularly polarized laser light. The laser polarization could be reversed. The electron polarization reached 77%, a number that made many electroweak parameters accessible to precise measurement. Electron polarization at the SLC produced right- or left-handed electrons at the point of intersection of the two beams. After interaction the electron polarization was measured by Compton back scattering of polarized laser light. The polarized Compton formula is discussed in Chapter 4. The positrons were not polarized, and this was also checked with the Compton polarimeter.

In order to understand the capability of the SLC to study electroweak physics we must insert polarization into Equation (7.114). First write the Z part in terms of right- and left-handed coefficients, restoring the fermion labels, and calling the muon "f" for an arbitrary fermion (lepton or quark) in the final state:

$$\begin{aligned}\langle f\bar{f}|\, T\, |e^-e^+\rangle = {} & \frac{-g^2}{4\cos^2\theta_W(s-M_Z^2+iM_Z\Gamma_Z)}\bar{v}(p_2)\gamma^\lambda(g_L^e(1-\gamma^5) \\ & + g_R^e(1+\gamma^5))u(p_1)\bar{u}(p_3)\gamma_\lambda(g_L^f(1-\gamma^5) \\ & + g_R^f(1+\gamma^5))v(p_4). \end{aligned} \tag{7.116}$$

Only the electron tensor is affected, unless (and this is in some cases possible) the polarizations of the final state fermions are measured. We will be content to work out the general formula for Z production by polarized electrons. Recall Equation (6.49) and the discussion of polarization in muon decay. In the completeness sum, make the substitution:

$$(\not{p}_1+m)\to(\not{p}_1+m)\frac{1+\lambda\gamma^5\not{s}_1}{2}; \quad s_1=(E,0,0,E)/m=p_1/m. \tag{7.117}$$

Here, we are in the Z boson center of mass, and the electron four momentum and its spin vector are the same, except for the electron mass $1/m$. We show in Chapter 6 that for $(V-A)\times(V-A)$ we may substitute in the final formulas:

$$p_1\to\frac{p_1-\lambda m s_1}{2}=\frac{p_1(1-\lambda)}{2}; \quad (V-A)\times(V-A). \tag{7.118}$$

The companion formula for $(V+A)\times(V+A)$ is left as an exercise:

$$p_1\to\frac{p_1+\lambda m s_1}{2}=\frac{p_1(1+\lambda)}{2}; \quad (V+A)\times(V+A). \tag{7.119}$$

Cross terms of the form $(V\pm A)\times(V\mp A)$ do not contribute to the trace. The parameter λ is the electron polarization, and $\lambda=+1$ is 100% positive helicity. The BW and other kinematic factors multiplying the cross-section are not needed in a discussion of polarized beam effects, so we will be content with the contraction of the electron tensor with the final state fermion tensor. First we write down $E^{\sigma\lambda}$ before substitution of the electron polarization vector:

$$\begin{aligned} E^{\sigma\lambda} = 8\{&(g_L^2 + g_R^2)(p_1^\sigma p_2^\lambda + p_1^\lambda p_2^\sigma - p_1 \cdot p_2 g^{\lambda\sigma}) \\ &- (g_L^2 - g_R^2) i\epsilon^{\alpha\lambda\beta\sigma} p_{2\alpha} p_{1\beta}\}. \end{aligned} \tag{7.120}$$

The fermion tensor is obtained by the substitution $p_1 \rightarrow p_4$, and $p_2 \rightarrow p_3$. Thus,

$$\begin{aligned} F_{\sigma\lambda} = 8\{&((g_L^f)^2 + (g_R^f)^2)(p_{4\sigma} p_{3\lambda} + p_{4\lambda} p_{3\sigma} - p_4 \cdot p_3 g_{\lambda\sigma}) \\ &- ((g_L^f)^2 - (g_R^f)^2) i\epsilon_{\nu\lambda\mu\sigma} p_3^\nu p_4^\mu\}. \end{aligned} \tag{7.121}$$

The unpolarized cross-section is proportional to the contraction of these two tensors:

$$\begin{aligned} E^{\sigma\lambda} F_{\sigma\lambda} = 256\{&g_L^2 (g_L^f)^2 (p_1 \cdot p_4)(p_2 \cdot p_3) + g_R^2 (g_R^f)^2)(p_1 \cdot p_4)(p_2 \cdot p_3) \\ &+ g_L^2 (g_R^f)^2 (p_1 \cdot p_3)(p_2 \cdot p_4) + g_R^2 (g_L^f)^2 (p_1 \cdot p_3)(p_2 \cdot p_4)\}. \end{aligned} \tag{7.122}$$

The angular dependence in Equation (7.113) is reproduced here, if one sets $g_L^f = g_L$ and $g_R^f = g_R$. The terms are written separately because when we insert the polarization, g_L and g_R get opposite signs. Defining θ as in Equation (7.113), between the incident electron and the outgoing fermion, inserting $p_1 \rightarrow p_1(1-\lambda)/2$ for g_L, and $p_1 \rightarrow p_1(1+\lambda)/2$ for g_R gives:

$$\begin{aligned} E^{\sigma\lambda} F_{\sigma\lambda} = 16s^2\{&g_L^2 (g_L^f)^2 (1-\lambda)(1+\cos\theta)^2 \\ &+ g_R^2 (g_R^f)^2 (1+\lambda)(1+\cos\theta)^2 \\ &+ g_L^2 (g_R^f)^2 (1-\lambda)(1-\cos\theta)^2 \\ &+ g_R^2 (g_L^f)^2 (1+\lambda)(1-\cos\theta)^2\}. \end{aligned} \tag{7.123}$$

Rearranging terms and factoring out $(g_L^2 + g_R^2)((g_L^f)^2 + (g_R^f)^2)$ gives the final form:

$$\begin{aligned} E^{\sigma\lambda} F_{\sigma\lambda} = 16s^2 (g_L^2 + g_R^2)((g_L^f)^2 + (g_R^f)^2)\{&(1 - \lambda A^e)(1 + \cos^2\theta) \\ &+ 2A^f (A^e - \lambda)\cos\theta\}, \end{aligned} \tag{7.124}$$

where the asymmetry parameter is defined by

$$A^f = \frac{(g_L^f)^2 - (g_R^f)^2}{(g_L^f)^2 + (g_R^f)^2} = \frac{1 - 8t_{3L}Qs_W}{1 - 8t_{3L}Qs_W + 8Q^2 s_W^2}, \tag{7.125}$$

and $\lambda = +1$ is 100% polarized electrons in the beam direction. Equation (7.124) is the defining equation for the polarization analysis of the Z boson by the SLC/SLD experiment at Stanford in the 1990s. The electron asymmetry parameter A^e can be measured simply by counting the number of Z's produced by $\lambda = +1$ and $\lambda = -1$. Since the electron polarization is less than one, use $0 \leq P_e \leq 1$ instead. The fact that $|P_e| \sim 75\%$ greatly improved the sensitivity to A^e. Integrate Equation (7.124) over the solid angle. Then measure the total cross-section for Z production as a function of the sign of P_e:

$$\begin{gathered} N_+ = N_{\text{tot}}(1 - P_e A_e)/2, \quad N_- = N_{\text{tot}}(1 + P_e A_e)/2, \\ \frac{N_+ - N_-}{N_{\text{tot}}} = -P_e A_e \pm \sqrt{N_{\text{tot}}}/N_{\text{tot}}, \end{gathered} \tag{7.126}$$

where we have appended a statistical error to the measurement. Note that if $P_e \sim 37\%$ instead of 75%, you have to run four times as long to get the same statistical precision. For the electron Equation (7.125) becomes:

$$A^e = \frac{1 - 4s_W}{1 - 4s_W + 8s_W^2}, \quad s_W = \sin^2\theta_W = 0.223(\text{PDG}), \quad A^e = 0.215. \tag{7.127}$$

The SLC/SLD measurement from the total data sample of several years running the linear collider is

$$A^e = 0.15138 \pm 0.000216, \quad \sin^2\theta_W = 0.23097 \pm 0.00027. \tag{7.128}$$

We mentioned earlier that the value of the Weinberg angle depended on the defining process, and that the variations were accounted for by radiative corrections to the first-order electroweak theory. Here is an example of this effect. $\sin^2\theta_W$ is about 5% larger than the PDG value based on $1 - M_W^2/M_Z^2$, which is rather large for a radiative correction in QED, but is typical of corrections to the first order electroweak theory. Note that while the measured value of the asymmetry parameter A^e is about 20% smaller than the value in Equation (7.127), $\sin^2\theta_W$ changed by only 5%.

A^e can be measured from the electron polarization dependence of the total Z production rate. It is a small quantity, so the presence of the electron polarization in the asymmetry term $2A^f(A^e - P_e)\cos\theta$ increases the sensitivity to A^f, the asymmetry parameters for the other fermions. From Table 7.2, we expect A^f to be large for both (u, c),and (d, s, b) quarks. Electroweak theory predicts the asymmetry parameter for $Z \to \nu\bar{\nu}$ to be $A^\nu = 1$, but, although as we shall see the branching ratio into neutrinos is large, they are 100% invisible. We do not directly observe the quarks either, but by identifying heavy flavors in the final state the asymmetry parameters for c, b, and s quarks have actually been measured. Electron polarization is not required for these asymmetries, so LEP has also measured the heavy flavor parameters, but we will quote the results from SLC/SLD.[a] The decay $Z \to b\bar{b}$ has the cleanest signature, because with SLD secondary vertexing and particle identification many events with the decay chain $B \to D \to K$ were reconstructed. $Z \to c\bar{c}$ has no B meson, but one has to worry about events where there was a B, but it was not identified, so obtaining a pure c quark sample is not so easy, and this difficulty is amplified for $Z \to s\bar{s}$, where both B and D mesons are absent. In the electroweak theory b quarks and s quarks should be the same, so this is a test of that prediction. The SLC/SLD results are

$$\begin{gathered} A^b = 0.916 \pm 0.021, \quad A^c = 0.670 \pm 0.027, \\ A^s = 0.895 \pm 0.066(\text{stat}) \pm 0.062(\text{sys}). \end{gathered} \tag{7.129}$$

These results are consistent with the entries in Table 7.2.

---

[a] Both e^- and e^+ beams at LEP were transversely polarized by synchrotron radiation. Electron depolarization was used in [16] to measure the energy of the storage ring. Typical polarization was 40% — not small, but not constant over the length of a store. It would have been difficult, but possible to orient the spin longitudinally for collisions at the LEP detectors, but this property was not exploited in Z studies.

7.9. Z Production in $p\bar{p}$ Collisions

The derivation of the parton model formula for $q\bar{q} \to Z$ follows the same analysis for the W^+ outlined in Section 7.4. There is no quark mixing for the Z coupling, but as we have already seen each type of fermion has its own vector and axial vector coupling constants. As before, it is convenient to work with right- and left-handed couplings. The matrix element for $u\bar{u} \to Z$ is

$$-i\langle Z|\,T\,|u\bar{u}\rangle = \frac{-ig}{2\cos\theta_W}\bar{v}(p_2)\gamma^\mu(g_L^u(1-\gamma^5) + g_R^u(1+\gamma^5)u(p_1))\epsilon_\mu^{*\lambda}. \tag{7.130}$$

Use the completeness rule for the massive vector boson polarization:

$$\sum_\lambda \epsilon_\mu^{\lambda *}\epsilon_\nu^\lambda = -g_{\mu\nu} + \frac{q_\mu q_\nu}{M_Z^2}, \tag{7.131}$$

where $q = p_1 + p_2$, and contract the helicity projection operators to obtain:

$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}} |\,\langle Z|\,T\,|u\bar{u}\rangle\,|^2 = \frac{\pi\alpha}{2\cos^2\theta_W \sin^2\theta_W}\Big\{&-\mathrm{Tr}(\not{p}_2\gamma^\mu((g_L^u)^2(1-\gamma^5)\\ &+ (g_R^u)^2(1+\gamma^5))\not{p}_1\gamma_\mu) + \frac{1}{M_Z^2}\mathrm{Tr}(\not{p}_2\not{q}((g_L^u)^2\\ &\times(1-\gamma^5) + (g_R^u)^2(1+\gamma^5))\not{p}_1\not{q})\Big\}.\end{aligned} \tag{7.132}$$

No term involving γ^5 contributes to the traces, and $p_1 \cdot p_2 = M_Z^2/2$, so the final result is

$$\frac{1}{4}\sum_{\text{spins}} |\,\langle Z|\,T\,|u\bar{u}\rangle\,|^2 = \frac{2\pi\alpha}{\cos^2\theta_W \sin^2\theta_W}M_Z^2((g_L^u)^2 + (g_R^u)^2). \tag{7.133}$$

Repeating the steps that led to Equation (7.61) for W production, we have the cross-section for $u\bar{u} \to Z$:

$$d\hat{\sigma}(u\bar{u} \to Z) = \frac{2\pi^2\alpha}{\cos^2\theta_W \sin^2\theta_W}((g_L^u)^2 + (g_R^u)^2)\delta(\hat{s} - M_Z^2). \tag{7.134}$$

Adopting Equation (7.62) for Z production gives:

$$\begin{aligned} d\sigma(p\bar{p} \to Z) = \frac{K}{3} \frac{2\pi^2 \alpha}{\cos^2\theta_W \sin^2\theta_W} \int_0^1 dx_1 \int_0^1 dx_2 \{(u(x_1)u(x_2) \\ + \bar{u}(x_1)\bar{u}(x_2))((g_L^u)^2 + (g_R^u)^2) + (d(x_1)d(x_2) \\ + \bar{d}(x_1)\bar{d}(x_2) + 2s(x_1)s(x_2))((g_L^d)^2 + (g_R^d)^2)\} \\ \times \delta(\hat{s} - M_Z^2). \end{aligned} \tag{7.135}$$

All of the PDF's are for the proton, and the factor of two on $s(x_1)s(x_2)$ takes care of the strange antiquarks, that have the same distributions. The K factor of about 1.2 is the same as for W production, and the 1/3 is a color factor, accounting for the fact that only three color combinations out of the nine available are color neutral. The right- and left-handed couplings are given as a function of the Weinberg angle $s_W = \sin^2\theta_W$ as

$$\begin{aligned} (g_L^u)^2 + (g_R^u)^2 &= \frac{1}{4} - \frac{2}{3}s_W + \frac{8}{9}s_W^2, \\ (g_L^d)^2 + (g_R^d)^2 &= \frac{1}{4} - \frac{1}{3}s_W + \frac{2}{9}s_W^2. \end{aligned} \tag{7.136}$$

Following the W analysis, we change variables $dx_1 dx_2 = d\hat{s}dy_Z/s$, and integrate over $\hat{s}$ and the δ function to obtain the differential cross-section as a function of y_Z, the Z rapidity.

$$\begin{aligned} \frac{d\sigma(p\bar{p} \to Z)}{dy_Z} = \frac{K}{3} \frac{2\pi^2 \alpha}{\cos^2\theta_W \sin^2\theta_W} \frac{x_1 x_2}{M_Z^2} \{(u(x_1)u(x_2) + \bar{u}(x_1)\bar{u}(x_2)) \\ \times ((g_L^u)^2 + (g_R^u)^2) + (d(x_1)d(x_2) + \bar{d}(x_1)\bar{d}(x_2) \\ + 2s(x_1)s(x_2))((g_L^d)^2 + (g_R^d)^2)\}. \end{aligned} \tag{7.137}$$

Given the Z rapidity, $x_1 = (M_Z/\sqrt{s})e^{y_Z}$, and $x_2 = (M_Z/\sqrt{s})e^{-y_Z}$. The Z's are polarized by different amounts by u and d quarks (see Problem 7.10) and the asymmetry for $Z \to \mu^-\mu^+$ is small, as we learned above, but the rule of thumb mentioned for W production and decay, that the lepton follows the quark, still holds: $u \to Z \to \mu^-$. The Z is neutral, as is the $p\bar{p}$ initial state, so there is no preference for y_Z to be in one direction or the other. The numerical factor in Equation (7.137) is

$$\frac{K}{3}\frac{2\pi^2\alpha}{\cos^2\theta_W\sin^2\theta_W}\frac{(g_L^u)^2+(g_R^u)^2}{M_Z^2}=2.4\text{ nb.} \tag{7.138}$$

The same factor for W production, Equation (7.64), was 15 nb. Tevatron measurements from CDF $p\bar{p}$ collisions at 1.96 TeV [19] are

$$\begin{aligned}\sigma\times\mathrm{Br}(p\bar{p}\to W\to l\nu_l) &= 2.749\pm 0.010(\mathrm{stat})\ \pm\ 0.053(\mathrm{sys})\\ &\quad\pm 0.115(\mathrm{lum})\text{ nb,}\end{aligned} \tag{7.139}$$

and

$$\sigma\times\mathrm{Br}(p\bar{p}\to Z\to l\bar{l}) = 254.9\pm 3.3(\mathrm{stat})\pm 4.6(\mathrm{sys})\pm 15.2(\mathrm{lum})\text{ pb.} \tag{7.140}$$

In both cases ‘lepton’ refers to e or μ, the average, not the sum. The W cross-section quoted in Equation (7.40) at 1.8 TeV is about 10% smaller. The cross-section ratio $\sigma(W\to l\nu)/\sigma(Z\to ll)=10.8$. As noted earlier, this factor of 10 comes roughly from two factors of three, one for σ, and the other for the branching ratio. The uncertainty in luminosity, about 5%, is quoted separately with the intent to improve the calibration of the integrated luminosity in the course of time.

7.10. *Z* Decay Modes

We give the Particle Data Group listings, the companion to Table 7.1 for the Z boson, and then work out the various partial widths using the electroweak theory. The total width is $\Gamma_Z = 2.4952\pm 0.0023$ GeV (Table 7.3).

The last two columns refer to Equation (7.143), with $\sin^2\theta_W = 0.213$, the value obtained from $Z\to\nu\bar{\nu}$. The PDG table does not list the u and d type quarks separately. The factor of three is for color. The percentages come from the ratio Γ_{ff}/Γ_Z, where Γ_Z is the sum of all of the calculated partial widths, 2.462 GeV. The calculated percentages from Equation (7.143) are very close to the experimental

Table 7.3: Branching ratios for $Z \to f\bar{f}$.

Name	Branching ratio (%)	Γ exp (MeV)	Br theory (%)	Γ theory (MeV)
e^-e^+	(3.3632 ± 0.0042)	83.3	3.4	84
$\mu^-\mu^+$	(3.3662 ± 0.0066)	83.3	3.4	84
$\tau^-\tau^+$	(3.3696 ± 0.0083)	84.0	3.4	84
Invisible	(20.000 ± 0.0055)	499	20	499
$3 \times u\bar{u}$			12	293
$3 \times d\bar{d}$			15	374
Hadrons	(69.911 ± 0.056)	1746	69	1711

measurements. The overall width is about 1% low, using a value for the Wienberg angle that gives the correct neutrino rates. The neutrino final state should be the least subject to corrections. This is another example of slightly different numbers for θ_W. 'Invisible' refers to neutrinos. How do you measure that in an e^-e^+ collider? Unfortunately, you do not know when an e^-e^+ pair has disappeared from the beams. So, if you make a Z and nothing appears in the detector, you do not know it was made. One trick is to run at a known energy above the resonance, and look for initial state radiation of a γ ray with energy $E_\gamma = \sqrt{s} - M_Z$. If you detect the γ, you know that the Z is there, so you can measure how often you don't see it. The initial state γ ray acts as a tag. The electroweak theory has three light neutrino families, and that is a fundamental quantity that can be measured by the 20% branching ratio, or from the total width of the Z peak that can be observed by an energy scan of the e^-e^+ collider through the resonance. The sum of the charged lepton widths is 252 MeV, half of the invisible width, so that the branching ratio for $Z \to \nu_e\bar{\nu}_e$ is twice the branching ratio for $Z \to e^-e^+$. The branching ratios to all three charged lepton flavors are the same, consistent with lepton universality that was also present in the W case. The dominant branching ratio is $Z \to$ hadrons, that is 20 times more numerous than $Z \to \mu^-\mu^+$. The resonance is very strong, so if you scan the energy region from 75 to 105 GeV in 1 GeV steps and count charged particle activity, the 30 nb cross-section for $Z \to$ hadrons will knock you over.

The derivation of the $Z \to f\bar{f}$ branching ratios follows the procedure used for the W. The four momenta are $p = p_3 + p_4$, and the matrix element is

$$-i \left\langle f\bar{f}\right| T \left|Z\right\rangle = -\frac{ig}{2\cos\theta_W}\bar{u}(p_3)\gamma^\mu(g_L^f(1-\gamma^5) + g_R^f(1+\gamma^5))v(p_4)\epsilon_\mu^\lambda. \tag{7.141}$$

The usual calculation gives the spin average squared:

$$\frac{1}{3}\sum_{\text{spins}} \left| \left\langle f\bar{f}\right| T \left|Z\right\rangle \right|^2 = \frac{8\pi\alpha}{3\sin^2\theta_W\cos^2\theta_W}((g_L^f)^2 + (g_R^f)^2)M_Z^2. \tag{7.142}$$

Repeating the steps to Equation (7.74), we have for the transition rates:

$$\begin{aligned}\Gamma(Z \to f\bar{f}) &= \frac{\alpha}{6\sin^2\theta_W\cos^2\theta_W}M_Z((g_L^f)^2 + (g_R^f)^2)\\ &= 661((g_L^f)^2 + (g_R^f)^2)\ \text{MeV}.\end{aligned} \tag{7.143}$$

The numerical values, using $g_L = t_3 - Q\sin^2\theta_W$ and $g_R = -Q\sin^2\theta_W$ with $\sin^2\theta_W = 0.213$, are given in the last column of Table 7.3. To obtain the number for the hadrons, multiply the sum of u and d quarks in Table 7.3 by two for $c\bar{c}$ and $s\bar{s}$, then add $b\bar{b} = d\bar{d}$. This procedure neglects the b quark mass. According to the theory, each light neutrino gets 166 MeV of width, about 7% of the total width of the Z. So, if there were four neutrinos instead of three, and everything else was the same, $\Gamma_Z = 2.66\,\text{GeV}$ instead of 2.5, which indicates the sensitivity of the width measurement to the number of neutrino families.

This has been a long chapter, and with all of the calculational detail it is easy to lose sight of the glory of the electroweak theory, but it is truly a remarkable accomplishment. Both the W and the Z are very robust particles. The pattern of complicated couplings is confirmed by experiment. Quarks are a fundamental component of the theory. The intermediate bosons are massive, but are easy to produce and to detect if you have enough energy available. The Z is an outstanding resonance at 91 GeV. It is clean and clear, and really stands out above the background. Fifty years ago neither intermediate boson had been discovered, and the weak neutral currents were not known to exist. But now, with the discovery of the Higgs, it seems that all of the parts to the picture are with us in glorious confirmation of the electroweak theory. The remarkable agreement between theory and experiment justifies a closer look at the origins of these ideas in Chapter 8.

7.11. Problems

7.1 Polarized Møller scattering. Refer to Section 4.9 and Figure 4.4 for the unpolarized analysis. Both initial spins must be longitudinally polarized, so we need two projection operators. For t channel exchange $p_1 \to p_3$ and $p_2 \to p_4$ we have the T matrix amplitude squared:

$$\sum_{\text{spins}} |A(t)|^2 = \frac{e^4}{4t^2}\mathrm{Tr}\{\not{p}_3\gamma^\mu(\not{p}_1 + m)(1 + \lambda_1\gamma^5\not{s}_1)\gamma^\nu\} \\ \times \mathrm{Tr}\{\not{p}_4\gamma_\mu(\not{p}_2 + m)(1 + \lambda_2\gamma^5\not{s}_2)\gamma_\nu\}. \tag{7.144}$$

Evaluate the traces to obtain

$$\sum_{\text{spins}} |A(t)|^2 = \frac{2e^4}{t^2}\{p_1 \cdot p_2 p_3 \cdot p_4 + p_1 \cdot p_4 p_2 \cdot p_3 \\ - m^2\lambda_1\lambda_2(s_1 \cdot s_2 p_3 \cdot p_4 - s_1 \cdot p_4 s_2 \cdot p_3)\}. \tag{7.145}$$

In the center of mass with $E \gg m$ we have the four vectors: $p_1 = (E, 0, 0, E) = ms_1$ and $p_2 = (E, 0, 0, -E) = ms_2$. Then obtain the expression in terms of the Mandelstam (s, t, u) variables:

$$\sum_{\text{spins}} |A(t)|^2 = \frac{e^4}{2t^2}\left\{s^2 + u^2 - \lambda_1\lambda_2(s^2 - u^2)\right\}. \tag{7.146}$$

$\sum |A(u)|^2$, for the crossed diagram, is easy. Just interchange t and u in Equation (7.146). The cross term, which gets a minus sign, is not so easy, so it is left for extra credit for those who feel comfortable by now with doing traces all day long. The answer is $e^4 s^2(1-\lambda_1\lambda_2)/tu$, with a plus sign. Write down the cross-section apart from phase space as a function of λ_1 and λ_2, each of which can be ±1, for four possibilities. Add them all up and you get the unpolarized cross-section (multiply the spin independent part by four). Show that the formula

$$\sum_{\text{spins}} |A(t)-A(u)|^2 = \frac{e^4}{2}\left\{\frac{s^2+u^2-\lambda_1\lambda_2(s^2-u^2)}{t^2}\right. \\ \left. + \frac{s^2+t^2-\lambda_1\lambda_2(s^2-t^2)}{u^2} + \frac{2s^2}{tu}(1-\lambda_1\lambda_2)\right\}. \tag{7.147}$$

becomes for $\theta = 90^\circ$ scattering angle in the center of mass:

$$\frac{\sigma(m_s=\pm 1)}{\sigma(m_s=0)} = \frac{8}{1}, \tag{7.148}$$

where the quantization axis is along the incident beam. This demonstrates the large spin dependent effect for Møller scattering. It has been used to measure the longitudinal polarization of high energy electron beams. See [20].

7.2 **Solar neutrinos.** This is a subject of fundamental importance. There have been several earthbound detectors studying solar neutrinos, but we will consider in some detail only three: Super-Kamiokande-IV, a large water detector in a mine in Japan, the Homestake neutrino detector of the late Ray Davis in Lead, South Dakota, and the Sudbury Neutrino Observatory in a mine in Canada. Begin with Super-Kamiokande-IV, because it is a large water Cherenkov detector that has enough sensitivity to observe $\nu_e + e^- \to \nu_e + e^-$. Solar neutrinos come from fusion reactions in the core of the sun. There are several different fusion chains, the basic one (that has the highest rate but the lowest energy neutrinos) is $p+p+p+p \to {}^4\mathrm{He} + 2\mathrm{e}^+ + 2\nu_\mathrm{e}$. Note that these are electron neutrinos, not antineutrinos. Antineutrinos come from nuclear reactors. Solar neutrino energies are low, from about 1 to 15 MeV at the upper end. The flux at the earth can be calculated from the standard solar model, and typical numbers are in the range 10^6 to $10^7 \mathrm{cm}^{-2}\mathrm{s}^{-1}$. The flux is directional. It is coming from the sun, and at night it goes through the earth to reach the detector. The fiducial volume of Super-K is 22,500 metric tons of water. Calculate the number of events per day $\nu_e + e^- \rightarrow \nu_e + e^-$ for a flux of $5 \times 10^6\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ 8 MeV electron neutrinos. Each water molecule has 10 electrons. How many events per day if the electron neutrino oscillated into a muon neutrino on the way from the sun to the earth? The cross-section for $\nu_e e^- \rightarrow \nu_e e^-$ is isotropic in the center of mass. Sketch the angular distribution of recoil electrons in the water. This distribution has been used by Super-K to point back to the sun. See [21]. They report a flux of $2.3 \times 10^6\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$.

7.3 Next is Ray Davis. He and a few collaborators operated a tank filled with 378,000 liters of C_2Cl_4, perchloroethylene, density 1.62 $\mathrm{gm/cm}^3$ in a gold mine in South Dakota, nearly one mile underground, for 25 years. This experiment started well before Super-K, but it comes second for us because its cross-section for neutrino detection is not so easy to calculate. It is a nuclear reaction, and you recall that we pointed out at the beginning of our discussion of weak interactions in Chapter 6 that nuclear beta decay lifetimes range from microseconds to millions of years with the same coupling constant G_F. The reaction, proposed for solar neutrino detection by Bruno Pontecorvo, is $\nu_\mathrm{e} + {}^{37}\mathrm{Cl} \rightarrow {}^{37}\mathrm{Ar} + e^-$. There are two stable isotopes of naturally occurring Chlorine: 25% ^{37}Cl and 75% ^{35}Cl. ^{37}Ar captures a K shell electron to return to ^{37}Cl with a lifetime of 50 days. Argon is a rare gas, so it does not form compounds with the C_2Cl_4, and can be separated out from the tank by using helium gas and other procedures. The threshold neutrino energy for the cross-section is 800 keV, so these neutrinos can be softer than the ones detected by Super-K. The cross-section for Pontecorvo's reaction is $\sigma({}^{37}\mathrm{Cl}(\nu_\mathrm{e}, \mathrm{e}^-){}^{37}\mathrm{Ar}) = 1.1 \times 10^{-42}\,\mathrm{cm}^2$ for 7 MeV neutrinos from ^{8}B. The flux of these neutrinos is expected to be $5 \times 10^6\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$, according to the solar model. First calculate the production rate, and then the equilibrium content of ^{37}Ar in the tank, taking the decay rate into account. Compare this number with the number of ^{37}Cl atoms in the tank. Convert the production rate into Solar Neutrino Units, where one SNU $= 10^{-36}$ captures per target atom per second. Note that the SNU is a transition rate, cross-section times flux. The final result from Homestake is capture rate $= 2.56 \pm 0.16 \pm 0.16$ SNU [22].

7.4 SNO in Canada. Both yields in Problems 7.2 and 7.3 are a factor of two to three below the expectations for the solar neutrino flux. This was known for years as the solar neutrino problem. Possibilities were that the solar model was incorrect, or that the experiments were wrong, or both. It is now generally accepted that theory and experiment are OK, and that the neutrinos themselves are at fault. The electron neutrino changes flavor on its trip to the earth, and muon or tau neutrinos do not scatter as much from electrons, nor do they make ^{37}Ar. The secret to discovering this is the neutral current scattering of neutrinos by nucleons, that does not distinguish between the neutrino flavors, and has a much larger cross-section than neutrino–lepton scattering. The Sudbury Neutrino Observatory was a large sphere 12 m in diameter filled with 1000 metric tons of pure heavy water (much smaller than Super-K, but D_2O rather than H_2O) borrowed from the nuclear power industry in Canada. By detecting neutrons from the breakup of deuterium via $\nu_x + d \to \nu_x + n + p$, a NC reaction, and comparing that rate to $\nu_e + d \to e^- + p + p$, a CC reaction, they were able to show that there were neutrinos from the sun that were not electron neutrinos. The deuterium served as both the neutron target and the free neutron detector via the capture $n + {}^2\mathrm{H} \to {}^3\mathrm{H} + \gamma$, where $E_\gamma = 6$ MeV. See [24] We do not have a satisfactory derivation for either the cross-section for $^{37}\mathrm{Cl} \to {}^{37}\mathrm{Ar}$, or for the deuteron break up. Parton model formulas like Equation (7.21) do not hold at 5 MeV neutrino energy. Deuteron breakup requires 2 MeV, and the recoil nucleons have kinetic energy ~1 MeV, hardly the asymptotic high energy limit. The cross-sections predicted by the parton model when extrapolated down to $E_\nu \sim 5$ MeV are too large by about a factor of ten. Accurate derivation of the required breakup formulas is a challenge for theorists. See [23] for the cross-sections:

$$\begin{aligned} \text{CC } \sigma(\nu_e d \to e^- pp)\ 5\text{ MeV } E_\nu &= 0.34 \times 10^{-42}\text{cm}^2; \\ 8\text{ MeV } E_\nu &= 1.44 \times 10^{-42}\text{cm}^2, \end{aligned} \tag{7.149}$$

and

$$\begin{aligned} \text{NC } \sigma(\nu_e d \to \nu_e np)\ 5\text{ MeV } E_\nu &= 0.1 \times 10^{-42}\text{cm}^2; \\ 8\text{ MeV } E_\nu &= 0.56 \times 10^{-42}\text{cm}^2. \end{aligned} \tag{7.150}$$

The CC cross-section is only for electron neutrinos, while the NC cross-section is the same for electron, muon or tau neutrinos. Note that the cross-sections do not scale $\sim E_\nu$ because of threshold effects, and that the ratio NC/CC $\sim 1/3$. Take the 8-MeV numbers, and calculate the number of CC and NC events per day for the solar flux given in Problem 7.2, assuming no neutrino flavor change. SNO quotes a detection efficiency for the NC events of 14%, and 576 ± 49 NC events detected for 306 days running. Compare this with your calculated NC yield. The quoted depleted CC flux $1.76 \pm 0.05 \pm 0.09 \times 10^6$ cm^{-2}s^{-1} together with the NC rate that is in agreement with the solar model was clear evidence for neutrino flavor changing. The work outlined in these three problems won a Nobel Prize in physics.

7.5 Work out the electron mass term for the NC scattering $\nu_\mu e^- \to \nu_\mu e^-$, given in the text after Equation (7.96), which with the extra term is

$$\frac{1}{2}\sum_{\text{spins}} |\left\langle \nu_\mu e^- \right| T \left| \nu_\mu e^- \right\rangle|^2 = 16G_F^2(g_L^2 s^2 + g_R^2 u^2 + 2m^2 g_L g_R t). \tag{7.151}$$

The Mandelstam variables keeping the electron mass satisfy $s + t + u = 2m^2$. Show that the dimensionless invariant variable $y = 1 - (p_2 \cdot p_3)/(p_1 \cdot p_2) = -t/(s - m^2)$.

7.6 Combined CC and NC cross-section for $\nu_e + e^- \to \nu_e + e^-$. Both W exchange and Z exchange are t channel, and we have written down both amplitudes: Equation (7.2) for charged current

$$\begin{aligned}&\left\langle \nu_e e^- \middle| T \middle| \nu_e e^- \right\rangle \\ &\quad = \frac{G_F}{\sqrt{2}} \{\bar{u}(p'_\nu)\gamma^\lambda(1-\gamma^5)u(p_e)\bar{u}(p'_e)\gamma_\lambda(1-\gamma^5)u(p_\nu)\},\end{aligned} \tag{7.152}$$

and Equation (7.92) for the neutral current, written in the same notation:

$$\begin{aligned}&\left\langle \nu_e e^- \middle| T \middle| \nu_e e^- \right\rangle \\ &\quad = \frac{G_F}{\sqrt{2}} \{\bar{u}(p'_\nu)\gamma^\mu(1-\gamma^5)u(p_\nu)\bar{u}(p'_e)\gamma_\mu(g_V - g_A\gamma^5)u(p_e)\}.\end{aligned} \tag{7.153}$$

Comparing these two formulas, apart from the different constants in the electron term for the neutral current, it is apparent that they have the same form if two of the spinors are interchanged, thus swap $\bar{u}(p'_\nu)$ and $\bar{u}(p'_e)$ in Equation (7.152). It is a feature of the algebra of the Dirac matrices that this swap can be done, with the price of a minus sign. In terms of components of the matrices:

$$(\gamma^\mu(1-\gamma^5))_{\alpha\beta}(\gamma_\mu(1-\gamma^5))_{\rho\sigma} = -(\gamma^\mu(1-\gamma^5))_{\rho\beta}(\gamma_\mu(1-\gamma^5))_{\alpha\sigma}. \tag{7.154}$$

This is called a Fierz transformation, after the theorist who first derived these formulas, and it is by no means obvious that you can do this. The transformation for $(V - A)$ is particularly simple, because it maps into $(V - A)$. The general reordering theorem involves all five invariants (S, V, T, A, P) and is a 5×5 matrix given in Appendix A of Barger and Phillips. Application of the reordering theorem to problems in Dirac algebra can simplify some calculations. So, we may rewrite Equation (7.152) as

$$\begin{aligned}&\left\langle\nu_e e^-\right| T \left|\nu_e e^-\right\rangle\\ &\quad= -\frac{G_F}{\sqrt{2}}\{\bar{u}(p'_e)\gamma^\lambda(1-\gamma^5)u(p_e)\bar{u}(p'_\nu)\gamma_\lambda(1-\gamma^5)u(p_\nu)\},\end{aligned} \tag{7.155}$$

There is a bit of subtlety regarding the relative algebraic sign of the two diagrams to begin with. Two fermion lines are interchanged between the W and Z exchange diagrams, and that gives a relative minus sign, that is canceled by the Fierz transformation. So, the final answer is

$$\begin{aligned}\left\langle\nu_e e^-\right| T \left|\nu_e e^-\right\rangle = \frac{G_F}{\sqrt{2}}&\left\{\bar{u}(p'_\nu)\gamma^\mu(1-\gamma^5)u(p_\nu)\bar{u}(p'_e)\gamma_\mu((1+g_V)\right.\\ &\left.-(1+g_A)\gamma^5)u(p_e)\right\}.\end{aligned} \tag{7.156}$$

Obtain the cross-section for $\nu_e e^-$ scattering for the sum of CC and NC exchange.

7.7 Take Equation (7.65) for the decay amplitude $W^+ \to \mu^+\nu_\mu$. Suppose the production mechanism was $u\bar{d} \to W^+$, so that the W^+ spin is along the $\bar{d}$ direction, opposite to the proton beam, with $m_s = -1$, or the polarization vector for the W is $\epsilon = (0, 1, -i, 0)/\sqrt{2}$. Evaluate the trace for this polarization vector, and show that the μ^+ angular distribution in the W rest frame has the form $(1-\cos\theta)^2$, where θ is the polar angle between the μ^+ and the proton beam. The μ^+'s, being right handed, are emitted backwards. Equation (7.36) is the solution to both production and decay. This exercise involves manipulating a spin one polarization vector. For the neutrino, which is the lepton in the final state, the sign of the $\cos\theta$ term flips, confirming the rule that the lepton follows the quark (u in the proton).

7.8 This problem requires access to a computer. Generate the angular distributions for μ^+ and μ^- relative to the proton beam direction for $p + \bar{p} \to W^\pm \to \mu^\pm + \nu_\mu$, for $\sqrt{s} = 2\,\text{TeV}$. Assume the parton model with no transverse momentum, so the $W^{\pm}$ are collinear with the beams. Calculate $y_{\text{max}} = \ln((E_{\text{max}} + p_{\text{max}})/(E_{\text{max}} - p_{\text{max}}))/2$ for the W's. Now generate y_W flat between -2 and $+2$, and shift the distribution by one unit towards the protons for W^+ and the antiprotons for W^-. This is an approximate model of the statement that W^+'s tend to follow protons, and W^-'s antiprotons. Hop on the W rest frame, and pick the muon decay angle according to the distribution $(1 \pm \cos\theta)^2$. Calculate the muon rapidity in the W rest frame, and use the fact that rapidities add to obtain separately the $\mu^{\pm}$ rapidities in the lab. The polar angle distributions then follow from the fact that the muon transverse momentum is Lorentz invariant. There are many ways to generate distributions with random numbers. Several dimensions require sophisticated techniques, but one dimension is easy, and there are at least two ways to do it. If $f(x)$ is normalized:

$$\int_a^b dx f(x) = 1,$$

and R is a random number $0 \le R \le 1$. Then let

$$R = \int_a^x dy f(y),$$

and solve for x. x will be distributed according to $f(x)$. Since $f(x)$ is frequently a quadratic equation, as is the case for these angular distributions, this leads to solutions to cubic equations. For $f(x) = 3(1+x)^2/8$ the solution is particularly simple: $x = 2R^{1/3} - 1$. If $f(x)$ is not convenient to integrate, there is another way. Plot $f(x)$ vs. x in the range of interest, and normalize it so that for $a \le x \le b$ $f(x)_{\text{max}} = 1$. Then pick $x_0 = (a-b)R_1 + b$, and compare $f(x_0)$ with a second random number R_2. If $f(x_0) > R_2$, the random number is inside the curve, keep the event, otherwise discard it and start over with a new value for R_1. This will also generate the curve, and is free of hard labor.

If you have access to PDF's, you can generate 'weighted events' by multiplying the flat W rapidity distribution with

$-y_{\max} \leq y_W \leq +y_{\max}$ by $x_1 x_2 u(x_1) d(x_2)$ for W^+ and $x_1 x_2 d(x_1) u(x_2)$ for W^-, where $x_{1,2} = M_W e^{\pm y_W}/\sqrt{s}$. If you do this the W rapidity distribution will have the correct shape, and will be shifted towards the protons for W^+ and towards the antiprotons for W^-. You will probably need 100 times as many events to get comparable numbers. This has no effect on the decay in the W center of mass.

7.9 Calculate the decay rate for $t \to W^+ + s$, starting from first principles, Equation (7.73). Use the appropriate CKM matrix element from Equation (7.70). What is the branching ratio for $\Gamma(t \to Ws)/\Gamma(t \to all)$? How would you search for this decay in a very large $t\bar{t}$ event sample?

7.10 The helicity rules do not hold for $p\bar{p} \to Z$ production, because of the right-handed quarks, but left-handed still prevails, and although for $Z \to \mu^+\mu^-$ the asymmetry is small, it still favors left-handed fermions, so the rule that the lepton follows the quark holds for both W's and Z's — just not as convincing for Z's. Take Equation (7.130) and leave the polarization vector $\epsilon_\mu^{*\lambda}$ in the calculation to repeat the analysis done for W's beginning with Equation (7.49). Remember to lower the indices for the quark four momenta! You should obtain an expression like:

$$\sum_{\text{quark spins}} |\langle Z| T |u\bar{u}\rangle|^2 = \frac{2g^2}{\cos^2\theta_W}\{(g_L^2 + g_R^2)(p_1 \cdot \epsilon^\lambda p_2 \cdot \epsilon^{\lambda *} + p_1 \cdot \epsilon^{\lambda *} p_2 \cdot \epsilon^\lambda + p_1 \cdot p_2) + (g_R^2 - g_L^2) i\epsilon^{\alpha\mu\beta\nu}\epsilon_\mu^{\lambda *}\epsilon_\nu^\lambda p_{2\alpha} p_{1\beta}\}. \tag{7.157}$$

Now use the polarization vector components defined by Equation (7.25) to obtain:

$$\sum_{\text{quark spins}} |\langle Z| T |u\bar{u}\rangle|^2 = \begin{cases} \dfrac{16G_F}{\sqrt{2}} g_R^2 M_Z^4, \ \lambda = +1; \\ \dfrac{16G_F}{\sqrt{2}} g_L^2 M_Z^4, \ \lambda = -1; \\ 0, \ \lambda = 0. \end{cases} \tag{7.158}$$

The sum of these two terms divided by four for initial spins agrees with Equation (7.133). Calculate the Z polarization for $u\bar{u}$ and $d\bar{d}$ production in $p\bar{p}$ collisions, using the numbers in Table 7.2.

7.11 Show that the angular distribution of the final state fermion f for $Z \to f\bar{f}$ with the Z polarization state $\lambda = -1$ in the Z rest frame is

$$\sum_{f\bar{f}\text{ spins}} |\langle f\bar{f}|\, T\, |Z\rangle|^2 = \frac{g^2 M_Z^2}{2\cos^2\theta_W}\Big\{(g_L^f)^2(1+\cos\theta)^2 + (g_R^f)^2(1-\cos\theta)^2\Big\}. \qquad (7.159)$$

θ is the polar angle of the outgoing fermion, and since the initial spin $m_s = -1$, the left-handed fermion likes to go forwards.

References

[1] S. Weinberg, *Phys. Rev. Lett.* **19**, 1264 (1967).
[2] D. Haidt, Discovery of weak neutral currents, in H. Schopper and L. DiLella (eds.), *60 Years of CERN Experiments and Discoveries* (World Scientific, Singapore, 2015).
[3] C. L. Cowan, Jr, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, *Science* **124**(3212) (1956).
[4] A. Burrows and J. M. Lattimer, *Astrophys. J.* **318**, 163 (1987).
[5] B. Barish *et al.*, *Phys. Rev. Lett.* **34**, 538 (1975).
[6] J. P. Berge *et al.*, *Zeitschrift für Physik C* **49**, 187 (1991).
[7] J. P. Berge *et al.*, *Zeitschrift für Physik C* **35**, 443 (1987).
[8] T. Affolder *et al.* (CDF Collaboration), *Phys. Rev. D* **3**, 072003 (2001).
[9] N. Cabibbo, *Phys. Rev. Lett.* **10**, 531 (1963).
[10] M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49**, 652 (1973).
[11] M. Tanabashi *et al.* (Particle Data Group), *Phys. Rev. D* **98**, 030001 (2018).
[12] L. Wolfenstein, *Phys. Rev. Lett.* **51**, 1945 (1983).
[13] T. Aaltonen *et al.* (CDF), *Phys. Rev. Lett.* **111**, 202001 (2012).
[14] T. Aaltonen *et al.* (CDF & D0), *Phys. Rev. Lett.* **115**, 152003 (2014).
[15] T. Aaltonen *et al.* (CDF & D0), *Phys. Rev. Lett.* **120**, 042001 (2018).
[16] R. Assmann *et al.*, *Zeitschrift für Physik C* **66**, 567 (1995).
[17] K. Abe *et al.* (SLD Collaboration), *Phys. Rev. Lett.* **84**, 5945 (2000).
[18] P. C. Rowson, D. Su, and S. Willocq, *Ann. Rev. Nucl. Part Sci.* **51**, 345–412 (2001).

[19] A. Abulencia *et al.* (CDF Collaboration), *J. Phys. G* **34**, 2457 (2007).
[20] H. Band *et al.*, *Nucl. Instr. Meth. A* **400**, 24 (1997).
[21] K. Abe *et al.* (Super-K-IV Collaboration), *Phys. Rev. D* **94**, 052010 (2016).
[22] B. T. Cleveland *et al.*, *Astrophys. J.* **496**, 505 (1998).
[23] M. Butler *et al.*, *Phys. Rev. C* **63**, 035501 (2001).
[24] Q. R. Ahmad *et al.*, Sudbury neutrino observatory, *Phys. Rev. Lett.* **89**, 011301 (2002).

This page intentionally left blank

Chapter 8

Electroweak Theory

8.1. Introduction

The electroweak theory of Glashow, Weinberg, and Salam (Nobel Prize 1979) that was analyzed in some detail in Chapter 7, and that agrees very well with experiment, is based on gauge invariance, patterned after electromagnetism. In this chapter, we will sketch the derivation of the electroweak theory to show where the intricate coupling rules discussed in Chapter 7 come from. As a bonus, we will also obtain the Higgs boson, and learn something about its couplings and decay modes. The discovery of the Higgs boson at the LHC caused quite a stir in 2012, and resulted in a Nobel Prize for Peter Higgs and co-discoverers in 2013.

Gauge invariance is important in quantum field theory, because it guarantees that the theory is renormalizable, a feature that is beyond the scope of this text. In a renormalizable theory, higher-order diagrams and radiative corrections can be treated in a way that preserves convergence, so that corrections to the first-order theory can be calculated. QED is a renormalizable theory based on invariance under a local gauge transformation of the vector potential:

$$A^\mu \to A^\mu + \partial^\mu \chi(x), \tag{8.1}$$

where $\chi(x)$ is an arbitrary function of the space–time variables. The discussion around Equation (3.204) outlines gauge transformations of the QED Lagrangian density for an electron and an electromagnetic field in terms of the field tensor $F^{\mu\nu}$:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu,$$

then

$$\mathcal{L} = -F_{\mu\nu}F^{\mu\nu}/4 + \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + eA_\mu \bar{\psi}\gamma^\mu \psi. \tag{8.2}$$

Introduce the covariant derivative:

$$D_\mu = \partial_\mu - ieA_\mu; \quad \mathcal{L} = -F_{\mu\nu}F^{\mu\nu}/4 + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi. \tag{8.3}$$

If the vector potential is changed by the gauge transformation Equation (8.1), the Dirac spinor must be changed by a phase to retain gauge invariance of the Lagrangian:

$$\psi(x) \to e^{ie\chi(x)}\psi(x), \quad D_\mu \to D_\mu - ie(\partial_\mu \chi), \tag{8.4}$$

and

$$(D_\mu - ie(\partial_\mu \chi))e^{ie\chi}\psi = e^{ie\chi}D_\mu \psi, \quad \bar{\psi} \to \bar{\psi}e^{-ie\chi}. \tag{8.5}$$

Hence $\mathcal{L}$ in Equation (8.3) is gauge invariant. This is a local gauge transformation, because χ is a function of x. In group theory language it is a U(1) transformation. The electron mass term is gauge invariant, but a photon mass term of the form $M^2 A^\mu A_\mu$ in the Lagrangian would not be, as discussed in Chapter 3. In deriving the Euler–Lagrange equations it is necessary to remember that all of the indices in $F_{\mu\nu}F^{\mu\nu}$ are summed over, and that is where the factor 1/4 comes in. It works like this. Lower all of the indices, and add in two metric tensors. Then be sure to take the functional derivative with respect to new dummy indices.

$$\begin{aligned}
-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} &= -\frac{1}{4}g^{\alpha\mu}g^{\beta\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial_\alpha A_\beta - \partial_\beta A_\alpha) \\
\frac{\partial \mathcal{L}}{\partial(\partial_\lambda A_\sigma)} &= -\frac{1}{4}\{g^{\alpha\lambda}g^{\beta\sigma}(\partial_\alpha A_\beta - \partial_\beta A_\alpha) - g^{\alpha\sigma}g^{\beta\lambda}(\partial_\alpha A_\beta - \partial_\beta A_\alpha) \\
&\quad + g^{\lambda\mu}g^{\sigma\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu) - g^{\sigma\mu}g^{\lambda\nu}(\partial_\mu A_\nu - \partial_\nu A_\mu)\} \\
&= \partial^\sigma A^\lambda - \partial^\lambda A^\sigma = F^{\sigma\lambda}.
\end{aligned} \tag{8.6}$$

For the free field Lagrangian $\partial\mathcal{L}/\partial A_\sigma = 0$, so the normalization does not matter, but here we have a source term from the covariant derivative:

$$\frac{\partial}{\partial x^\lambda}\frac{\partial\mathcal{L}}{\partial(\partial_\lambda A_\sigma)} = \partial^\sigma\partial_\lambda A^\lambda - \partial_\lambda\partial^\lambda A^\sigma = \partial_\lambda F^{\sigma\lambda} = \frac{\partial\mathcal{L}}{\partial A_\sigma} = e\bar{\psi}\gamma^\sigma\psi. \tag{8.7}$$

It is worth repeating the classical Equation (3.90) for comparison:

$$\vec{\nabla}(\vec{\nabla}\cdot\vec{A} + (1/c^2)\partial V/\partial t) - \vec{\nabla}^2\vec{A} + (1/c^2)\partial^2\vec{A}/\partial t^2 = \mu_0\vec{j}. \tag{8.8}$$

Apart from the units and the compressed notation, the space components of Equations (8.7) and (8.8) are exactly the same, with the space part of the Dirac four current $\vec{j} = e\bar{\psi}\vec{\gamma}\psi$. There is no mass term in the wave equation. Since Equation (8.6) is antisymmetric in σ and λ, the current density is conserved:

$$\partial_\sigma\partial^\sigma\partial_\lambda A^\lambda - \partial_\sigma\partial_\lambda\partial^\lambda A^\sigma = e\partial_\sigma\bar{\psi}\gamma^\sigma\psi = 0. \tag{8.9}$$

8.2. Non-Abelian Gauge Transformations

The concept of weak I-spin was introduced in Section 6.6 of Chapter 6. For leptons and quarks, left-handed are assigned weak I-spin 1/2, and right-handed weak I-spin 0, and the $W^\pm$ bosons weak I-spin 1. There are no right-handed neutrinos. The right-handed leptons and quarks do not couple to the W's, although as we have studied in Chapter 7, they do couple to the Z. Weak I-spin is incorporated into the electroweak theory as a non-abelian gauge transformation.

We begin our discussion with Dirac spinors ψ that are two component vectors in an abstract I-spin space, and the gauge transformation is a rotation in that space:

$$\psi' = e^{-ig\chi_i T_i}\psi, \tag{8.10}$$

where g is a constant, and χ_i has three components. The hermitian operators T_i satisfy $[T_i, T_j] = i\epsilon_{ijk}T_k$, and for I-spin 1/2 $T_i = \sigma_i/2$, σ_i being the usual Pauli spin matrix. For an abelian gauge transformation, it is easy to use the exponential form, as we have done above, but for the non-abelian case, because of the complications involved with non-commuting operators in exponentials, it is simpler to work with the expansion to first order:

$$\begin{aligned} \psi' &= (1 - ig\chi_i T_i + O(\chi_i^2))\psi; \\ \bar{\psi}' &= \bar{\psi}(1 + ig\chi_j T_j + \ldots); \\ \bar{\psi}'\psi' &= \bar{\psi}\psi \text{ to first order in } \chi_i. \end{aligned} \tag{8.11}$$

The vector potential A_μ is replaced by a three-component vector in our abstract I-spin space:

$$W_{i\mu}, \quad i = 1, 2, 3. \tag{8.12}$$

The gauge transformation is

$$W_{i\mu} \to W_{i\mu} + \partial_\mu \chi_i + g\epsilon_{ijk}\chi_j W_{k\mu}. \tag{8.13}$$

The free field Lagrangian is

$$\mathcal{L} = -W_{i\mu\nu}W_i^{\mu\nu}/4, \quad W_{i\mu\nu} = \partial_\mu W_{i\nu} - \partial_\nu W_{i\mu} - g\epsilon_{ijk}W_{j\mu}W_{k\nu}. \tag{8.14}$$

The extra terms are required by the non-abelian gauge transformation. The ϵ terms, while a cross product in the abstract space, are also the 'structure constants' of the rotation group, as they appear in the commutation relations for the T_i's. Thus, while Equation (8.13) does not involve the T_i operators explicitly, the group presence is reflected in ϵ. We have one more relation before we attempt to justify these equations, and that is the covariant derivative that contains the T_i operators:

$$D_\mu = \partial_\mu + igW_{i\mu}T_i. \tag{8.15}$$

The commutator of two covariant derivatives, the analog of Equation (3.207) for electromagnetism, is

$$[D_\mu, D_\nu]\psi = igW_{k\mu\nu}T_k\psi, \tag{8.16}$$

and, as in Equation (3.207), the commutator is free of derivative operators, although it does involve the SU(2) operators T_k, because of the non-abelian nature of the gauge group. The Lagrangian may be written as

$$\mathcal{L} = -W_{i\mu\nu}W_i^{\mu\nu}/4 + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi, \tag{8.17}$$

This looks very much like Equation (8.3) above, but it is really a lot more complicated. Let us look at the Dirac part first. We have already established that the mass term is gauge invariant. We need to verify the relation to first order in χ:

$$D'_\mu(1 - ig\chi_i T_i) = (1 - ig\chi_i T_i)D_\mu,$$

where

$$D'_\mu = \partial_\mu + ig(W_{i\mu} + \partial_\mu \chi_i + g\epsilon_{ijk}\chi_j W_{k\mu})T_i. \tag{8.18}$$

So,

$$\begin{aligned}\partial_\mu + ig(W_{i\mu} + \partial_\mu\chi_i + g\epsilon_{ijk}\chi_j W_{k\mu})T_i - ig(\partial_\mu\chi_i)T_i - ig\chi_i T_i\partial_\mu \\ + g^2(W_{i\mu}T_i\chi_k T_k + i\epsilon_{ijk}\chi_j W_{k\mu}T_i) \\ = \partial_\mu + igW_{i\mu}T_i - ig\chi_k T_k\partial_\mu + g^2\chi_k T_k W_{i\mu}T_i.\end{aligned} \tag{8.19}$$

There is a lot of cancellation. One must be careful of the order of the non-commuting operators T_k. We are left with some terms proportional to g^2:

$$g^2(W_{i\mu}\chi_k(T_iT_k - T_kT_i) - i\epsilon_{kjl}\chi_j W_{k\mu}T_l) = 0. \tag{8.20}$$

This expression does indeed vanish, because of the commutation rules for the T's. Proof of the gauge invariance of the free field Lagrangian Equation (8.14) is more complicated, because the tensor $W_{i\mu\nu}$ is not gauge invariant. Only the product $W_{i\mu\nu}W_i^{\mu\nu}$ is invariant. Barger and Phillips [p. 39] assert that:

$$W_{i\mu\nu} \to W_{i\mu\nu} + g\epsilon_{ijk}\chi_j W_{k\mu\nu}. \tag{8.21}$$

If true, then the Lagrangian term transforms as

$$W_i^{\mu\nu}W_{i\mu\nu} \to W_i^{\mu\nu}W_{i\mu\nu} + g\epsilon_{ijk}\chi_j W_{k\mu\nu}W_i^{\mu\nu} + g\epsilon_{ilm}\chi_l W_m^{\mu\nu}W_{i\mu\nu}. \tag{8.22}$$

It is easy to show by interchanging the indices on one of the ϵ's that the terms proportional to g in Equation (8.22) cancel, leaving the Lagrangian gauge invariant. So, all we have to do is verify Equation (8.21). Writing Equation (8.21) in terms of W_μ gives two terms proportional to the gauge function χ:

$$\begin{aligned} W_{i\mu\nu} \rightarrow W_{i\mu\nu} &+ g\epsilon_{ijk}\chi_j(\partial_\mu W_{k\nu} - \partial_\nu W_{k\mu}) \\ &+ g^2(\vec{\chi}\cdot\vec{W}_\mu W_{i\nu} - W_{i\mu}\vec{\chi}\cdot\vec{W}_\nu). \end{aligned} \tag{8.23}$$

The dot products are of vectors in I-spin space. Substitute Equation (8.13) for the gauge transformed $W_{i\mu}$ into Equation (8.14) for $W_{i\mu\nu}$, and retain only terms linear in χ. After some algebra, being careful how you handle the indices, and using the identity:

$$\begin{aligned} \epsilon_{ijk}\epsilon_{ilm} &= \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}, \\ \epsilon_{123} &= +1, \\ \text{interchange any two} &= -1, \\ \text{any two the same} &= 0, \end{aligned} \tag{8.24}$$

you will find that Equation (8.21) is correct. This is a non-abelian gauge transformation, named after Yang and Mills [1], who were the first ones to work it out in the context of ordinary I-spin, and the Lagrangian Equation (8.17) is gauge invariant. The gauge transformation is a rotation in an abstract 3D space, where the rotation operators do not commute with each other, and this leads to complications in the covariant derivative and the definition of the energy tensor for the vector field. Our vector $W_{i\mu}$ has four components in space–time, and three components in the abstract I-spin space.

Equation (15.40) in Peskin and Schroeder, in our notation and with our sign conventions, is a formula for the Yang–Mills source wave equation, the analog of Equation (8.7) for the abelian theory. It reads

$$-\partial_\nu W_i^{\nu\sigma} - g\epsilon_{ijk}W_{j\nu}W_k^{\nu\sigma} + g\bar{\psi}\gamma^\sigma T_i\psi = 0. \tag{8.25}$$

The first term, $\partial_\nu W_i^{\nu\sigma}$, is the analog of Equation (8.7). The rest of the field terms represent nonlinearities in the wave equation. The current density comes directly from the covariant derivative in the Lagrangian. The constant g is the analog of the elementary charge, and the operator T_i can change the T_3 quantum number of the two component wave function ψ. Like Equation (8.7), there is no mass term, which would be linear in $W_{i\nu}$. The Euler–Lagrange equations with $\mathcal{L}$ given by Equation (8.16) are

$$\frac{\partial}{\partial x^\nu}\frac{\partial \mathcal{L}}{\partial(\partial_\nu W_{i\sigma})} - \frac{\partial \mathcal{L}}{\partial(W_{i\sigma})} = 0. \tag{8.26}$$

Evaluating this equation requires considerably more labor than the abelian QED Lagrangian Equations (8.6) and (8.7). The factor of 1/4 is handled in the same way, but there are lots of terms, and you have the group index 'i' to worry about. Be extra careful about summed dummy indices. My answer for the first term is

$$\partial^\sigma\partial_\nu W_i^\nu - \partial_\nu\partial^\nu W_i^\sigma + g\epsilon_{ijk}((\partial_\nu W_j^\nu)W_k^\sigma + W_j^\nu\partial_\nu W_k^\sigma).$$

This is also the first term in Equation (8.25). The second part of the Euler–Lagrange equations is

$$\begin{aligned} &+g\epsilon_{ijk}W_{j\nu}(\partial^\sigma W_k^\nu - \partial^\nu W_k^\sigma) - g^2 W_{k\nu}(W_k^\nu W_i^\sigma - W_i^\nu W_k^\sigma) \\ &\quad + g\bar{\psi}\gamma^\sigma T_i\psi = 0, \end{aligned} \tag{8.27}$$

and these are the last two parts of Equation (8.25). The form for the Dirac source current looks perfectly ordinary, but it is not conserved. The divergence from Equation (8.25), since $\partial_\sigma\partial_\nu W_i^{\nu\sigma} = 0$, is simply:

$$\partial_\sigma g\bar{\psi}\gamma^\sigma T_i\psi = g\epsilon_{ijk}\partial_\sigma(W_{j\nu}W_k^{\nu\sigma}), \tag{8.28}$$

and the right-hand side does not vanish. This point is discussed by Yang and Mills in their original paper, and the interested reader is referred to that source. They point out that the reason for the nonzero Equation (8.28) is the fact that the $W_{k\nu}$ field carries I-spin, which is the analog of the charge, so that the Dirac current alone is not conserved. Moving the right-hand side of Equation (8.28) over to the left-hand side gives two terms, one from the fermions and one from the gauge bosons, and a conserved I-spin current. The same reason, that the gauge bosons are 'charged', accounts for the nonlinear terms in the $W_{k\nu}$ wave equation. The Dirac current of Yang and Mills is the nucleon current, with the I-spin operator T_+ changing a neutron into a proton, together with absorption of a W^+ or emission of a W^- vector boson. In 1954, the pion triplet π^+, π^0, π^- was known to couple to nucleons, but pions are pseudoscalars. The authors did not predict a mass for their vector boson. The first evidence for a strongly interacting vector meson, the ρ at 770 MeV mass, was reported in 1961 [2], seven years after the non-abelian I-spin model of Yang and Mills.

The original work of Yang and Mills did not have obvious applications at that time, but non-abelian gauge theories are at the core of the Standard Model. Modern QCD, beyond the scope of this text, is such a theory. The gauge group is color SU(3), so in the various formulas instead of ϵ_{ijk} one finds the f_{ijk} structure constants of the larger group. There are three colors, and each gauge boson is a color–anticolor pair, which would give nine gluons, but one is colorless, leaving eight massless gluons. This is the same group as discussed in Chapter 1, and the eight gluons form a color octet like pseudoscalar meson octet, with the three colors playing the role of the three original quarks. Our Yang–Mills gauge field $W_{i\mu}$ is a triplet in I-spin space, which is the SU(2) analog of the octet. In group theory language, the first combination of the smallest non-trivial representation of SU(2) is $2 \bigotimes 2 = 3 \bigoplus 1$, and of SU(3) is $3 \bigotimes 3^* = 8 \bigoplus 1$. QCD is also nonlinear, with cubic and quardic couplings of gluons with themselves, similar to Equation (8.27).

This chapter is about electroweak theory, for which we have an SU(2) symmetry group called weak I-spin, the same group first studied by Yang and Mills. But the gauge bosons have to be very massive to give the low energy theory that we studied in Chapter 6. So, in order to apply non-abelian gauge theory to the weak interactions, we have to find a way around the no mass problem. Sidney Coleman [3] attributes the 're-discovery' of Yang–Mills for electroweak theory to Steven Weinberg. But before we learn about the solution to the mass problem, called the Higgs mechanism, we have to assemble all of the parts that go into the theory.

8.3. The SU(2) × U(1) Lagrangian

We return to the left- and right-handed fermions and the weak I-spin model. It is worth repeating as we get started that weak I-spin is totally different from ordinary I-spin, except that they are the same group. Ordinary I-spin is a symmetry that is respected by the strong interactions, with a charge rule:

$$Q = I_z + \frac{Y}{2}; \quad Y = N + S. \tag{8.29}$$

The hypercharge $Y = N + S$ is the sum of the nucleon number N and the strangeness S, and is conserved by strong and electromagnetic interactions. Recall that in the $K^0\bar{K}^0/K_LK_S$ complex the hypercharge plays a role for the neutral K's analogous to the charge, hence the name. The charge rule is designed to account for the fact that the baryons with no strangeness have half integer I-spin, and the mesons integer I-spin, while the strange particles have I-spins interchanged. Only the weak interaction violates the conservation of both I_z and Y, preserving Q. The same names, weak I-spin and weak hypercharge, are used in the electroweak theory, but they do not have the same meaning. Weak I-spin is conserved by $W^\pm$ bosons and left-handed spin 1/2 Dirac fermions, and weak hypercharge is introduced to include the electric charge in a unified model.

The three component real vector field $W_{k\mu}$ introduced in the previous section now becomes a quantized field operator, with the following definitions:

$$W_+^\mu = \frac{W_1^\mu - iW_2^\mu}{\sqrt{2}}, \quad W_-^\mu = \frac{W_1^\mu + iW_2^\mu}{\sqrt{2}} = W_+^{\mu\dagger}, \quad W_0^\mu = W_3^\mu. \tag{8.30}$$

The sign convention chosen is opposite to that for the weak I-spin raising and lowering operators, thus,

$$T_\pm = \frac{T_1 \pm iT_2}{\sqrt{2}}; \quad \vec{W} \cdot \vec{T} = W_+T_+ + W_-T_- + W_0T_3. \tag{8.31}$$

T_+ acting on an electron turns it into a neutrino, so the field operator W_+ must create a negative W boson, and destroy a positive W boson.

For the fermions the left-handed states are assigned weak I-spin 1/2, and the right-handed states weak I-spin zero. There are no right-handed neutrinos in the model. The experimental evidence for neutrino flavor oscillations on the path from the sun to the earth, discussed in Problems 7.2–7.4, has been confirmed by several other experiments, and is clear evidence that the neutrinos have a very tiny mass, but not zero. Our model without right-handed neutrinos is therefore lacking, and has to be fixed, but for now we will ignore this problem. The definitions of left- and right-handed spinors are

$$\begin{aligned}&\psi_L = \frac{1-\gamma^5}{2}\psi; \quad \psi_R = \frac{1+\gamma^5}{2}\psi; \quad \psi = \psi_L + \psi_R;\\ &\bar{\psi}\psi = \bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L.\end{aligned} \tag{8.32}$$

The last relation follows because $\bar{\psi}_L\psi_L = \bar{\psi}_R\psi_R = 0$, and that means that there can be no fermion mass terms in a left-handed Lagrangian. Like the gauge fields, the fermions start out massless (Table 8.1).

The primes on the charge $-1/3$ quarks (d', s', b') imply mixing via the CKM matrix discussed in Chapter 7. The right-handed fermions assigned weak I-spin 0 are

$$\psi_R : (e^-, \mu^-, \tau^-, u, c, t, d, s, b)\ T = 0.$$

All quantum numbers are reversed for the antiparticles e^+, μ^+, etc.

The problems associated with identifying W_0 as the weak neutral current intermediate vector boson have already been discussed. Perhaps the most important reason is that it would be too restrictive, because the W's couple only to left-handed fermions. There is also the possibility of linking weak interactions and electromagnetism, for which we should include the electric charge somewhere in our formulas. To do this we introduce a fourth field, called B_μ, that is neutral, and invariant under abelian gauge transformations like the vector potential A_μ. Like everything else, B_μ is massless. We have a coupling constant g for the W field, so we will call the B field coupling constant g'.

Table 8.1: Left-handed weak I-spin fermion doublets.

Leptons			Quarks			T_3
ν_e	ν_μ	ν_τ	u	c	t	$+1/2$
e^-	μ^-	τ^-	d'	s'	b'	$-1/2$

Weak hypercharge is introduced as a parameter to allow the coupling of B_μ to the various particles to change. The electric charge appears in the theory through the charge rule:

$$Q = T_3 + \frac{Y_W}{2}. \tag{8.33}$$

This equation defines the weak hypercharge to look just like Equation (8.29), since both Q and T_3 are known for all of the leptons and quarks. The numbers are summarized in Table 8.2.

The same assignment obtains for the two other generations: $\nu_\mu, \mu^-, \nu_\tau, \tau^-, c, s, t, b$. The left-handed charge $-1/3$ quarks mix according to the CKM matrix. Y_W is reasonable for the left-handed fermions. You recall that the quarks have baryon number $N = 1/3$, so $Y_W = 1/3$ is perhaps not so peculiar. The Higgs scalar boson is a creature of the electroweak theory with carefully selected quantum numbers. The Z boson has mixed weak I-spin, part $T = 1$, and part $T = 0$. For the right-handed fermions Y_W is chosen to give the correct charge. The Lagrangian with W and B fields is

$$\mathcal{L} = -\frac{1}{4}W_i^{\mu\nu}W_{i\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} + \bar{\psi}i\gamma^\mu D_\mu\psi, \tag{8.34}$$

Table 8.2: $Q = T_3 + Y_W/2$.

ψ_L	T	T_3	Y_W	Q
ν_e	$1/2$	$+1/2$	-1	0
e^-	$1/2$	$-1/2$	-1	-1
u	$1/2$	$+1/2$	$1/3$	$+2/3$
d	$1/2$	$-1/2$	$1/3$	$-1/3$
ψ_R				
e^-	0	0	-2	-1
u	0	0	$+4/3$	$+2/3$
d	0	0	$-2/3$	$-1/3$
Higgs	$1/2$	$-1/2$	1	0
$W^\pm$	1	± 1	0	± 1
Z	Mixed	0	0	0

with

$$\begin{aligned} W_i^{\mu\nu} &= \partial^\mu W_i^\nu - \partial^\nu W_i^\mu - g\epsilon_{ijk} W_j^\mu W_k^\nu, \\ B^{\mu\nu} &= \partial^\mu B^\nu - \partial^\nu B^\mu, \\ D_\mu &= \partial_\mu + igW_{i\mu}T_i + ig' B_\mu \frac{Y}{2}. \end{aligned} \tag{8.35}$$

$W_i^{\mu\nu}$ is the same as Equation (8.14), and $B^{\mu\nu}$ is the same as Equation (8.2) for the vector potential. This covariant derivative serves for both left- and right-handed fermions, ψ_L and ψ_R, because the T_i term vanishes operating on right-handed fermions, that have $T = 0$, leaving only the B_μ part for ψ_R.

The Lagrangian Equation (8.34) is invariant under two independent gauge transformations, denoted as SU(2) × U(1). One is analogous to the electromagnetic vector potential, Equation (8.1), and the other was the subject of the previous section, Equation (8.13):

$U(1)_Y$	$SU(2)_L$
$B_\mu \to B_\mu + \partial_\mu \omega(x)$	$W_{i\mu} \to W_{i\mu} + \partial_\mu \chi_i(x) + g\epsilon_{ijk}\chi_j W_{k\mu}$
$\psi_L \to (1 - ig'\frac{Y}{2}\omega(x))\psi_L$	$\psi_L \to (1 - igT_i\chi_i(x))\psi_L$
$\psi_R \to (1 - ig'\frac{Y}{2}\omega(x))\psi_R$	$\psi_R \to \psi_R$
$W_{i\mu} \to W_{i\mu}$	$B_\mu \to B_\mu$

ψ_L and ψ_R transform the same under $U(1)_Y$, while only ψ_L is transformed under $SU(2)_L$. The Dirac field transformations come from the covariant derivative. Now expand the covariant derivative Equation (8.35), using Equation (8.31):

$$D_\mu = \partial_\mu + ig(W_\mu^+ T^+ + W_\mu^- T^- + W_\mu^{(3)} T_3) + ig' B_\mu \frac{Y}{2},$$

and isolate the neutral part:

$$igW_\mu^{(3)} T_3 + ig' B_\mu \frac{Y}{2}. \tag{8.36}$$

The weak hypercharge Y was introduced to couple with T_3 and bring the electric charge $Q = T_3 + Y/2$ into the formulas to link the weak and electromagnetic forces. As you can see from Equation (8.36),

T_3 and Y are separately mixed with the fields $W^{(3)}$ and B in the covariant derivative. New massless fields Z and A are introduced through a rotation matrix:

$$\begin{pmatrix} W^{(3)} \\ B \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z \\ A \end{pmatrix}. \tag{8.37}$$

The mixing angle θ_W is a free parameter of the theory that must be determined by experiment. Note that if $\theta_W = 0$ the rotation matrix is the unit matrix, with $W^{(3)} = Z$ and $B = A$. This choice leaves the neutral current couplings the same as the charged current couplings, and the electromagnetic current remains independent, meaning no electroweak unification. As we learned in Chapter 7, $\theta_W = 28.6°$, so the decoupled solution is not an option. Equation (8.37) reproduces Equation (7.86) for the mixed fields:

$$W_\mu^{(3)} = \cos\theta_W Z_\mu + \sin\theta_W A_\mu; \quad B_\mu = -\sin\theta_W Z_\mu + \cos\theta_W A_\mu. \tag{8.38}$$

Substitution of Equation (8.38) into Equation (8.36) gives:

$$\begin{aligned} igW_\mu^{(3)} T^3 + ig' B_\mu \frac{Y}{2} &= i\left(g\cos\theta_W T_3 - g'\sin\theta_W \frac{Y}{2}\right) Z_\mu \\ &\quad + i\left(g\sin\theta_W T_3 + g'\cos\theta_W \frac{Y}{2}\right) A_\mu. \end{aligned} \tag{8.39}$$

For the coefficient of A_μ to be the electric charge, we must have:

$$g\sin\theta_W = g'\cos\theta_W = e, \quad \tan\theta_W = \frac{g'}{g}, \tag{8.40}$$

where e is the charge of the positron. These relations among coupling constants clearly show that $\theta_W = 0$ is not an allowed value for the mixing angle. The hypercharge may be eliminated from the covariant derivative to leave:

$$igW_\mu^{(3)} T_3 + ig' B_\mu \frac{Y}{2} = i\frac{g}{\cos\theta_W}(T_3 - \sin^2\theta_W Q) Z_\mu + ieQA_\mu, \tag{8.41}$$

where Q is the charge of the Dirac spin 1/2 particle in units of the elementary charge. This expression is correct for $\psi_L = (1-\gamma^5)\psi/2$.

The T_3 term is absent for $\psi_R = (1+\gamma^5)\psi/2$. The current densities are

$$j_{\mu Z} = \frac{g}{\cos\theta_W}(\bar{\psi}_L\gamma_\mu(T_3 - \sin^2\theta_W Q)\psi_L - \bar{\psi}_R\gamma_\mu \sin^2\theta_W Q\psi_R), \tag{8.42}$$

and

$$j_{\mu A} = e(\bar{\psi}_L\gamma_\mu Q\psi_L + \bar{\psi}_R\gamma_\mu Q\psi_R). \tag{8.43}$$

Substituting in the helicity projection operators gives:

$$j_{\mu Z} = \frac{g}{2\cos\theta_W}\bar{\psi}\gamma_\mu(T_3 - 2\sin^2\theta_W Q - \gamma^5 T_3)\psi, \quad j_{\mu A} = e\bar{\psi}\gamma_\mu Q\psi. \tag{8.44}$$

Apart from the different notation $t_3 \to T_3$, the Z coupling is identical to Equations (7.87) and (7.90). This derivation gives the electroweak theory for left- and right-handed coupling to the neutral weak intermediate boson Z, and the relation between the coupling constant g and the elementary charge e. The consequences of this analysis have been confirmed by experiment in many different ways. The significant hypothesis of the electroweak theory is that the right-handed fermion coupling to the neutral weak intermediate boson Z can be supplied by electromagnetism.

The gauge fields and the Dirac fermions are still massless. In order to generate the masses we must invoke what is called the Higgs mechanism, and that technique is described in Section 8.4. It is fair to say that the Higgs mechanism is unfamiliar at best, but dismissing it without due consideration loses useful formulas like $M_W = M_Z\cos\theta_W$, and the basis for the Higgs scalar boson itself.

8.4. The Higgs Mechanism

In order to demonstrate spontaneous symmetry breaking it is necessary to introduce another scalar field to the Lagrangian. The simplest scalar field Lagrangian gives the Klein–Gordon equation, and was introduced in Chapter 3, Equations (3.26)–(3.29). An interaction term $-\lambda\phi^4/4$ was added to the Lagrangian for Equation (3.30), although we did not pursue the so-called ‘lambda phi fourth’ model any further. Since we will be manipulating that Lagrangian, it is useful to repeat it:

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4}\phi^4. \tag{8.45}$$

Note that the sign on the mass term gives the correct sign in the Klein–Gordon equation. If $\lambda = 0$ the Euler–Lagrange equations are $\partial_\mu \partial^\mu \phi + m^2\phi = 0$, or $-\omega^2 + \vec{k}^2 + m^2 = 0$. Also note the units. $\mathcal{L}$ is a Lagrangian density, with units E/l^3, but in our natural units $E \sim 1/l$, so the units are E^4, meaning $\phi^2 \sim E^2$, and λ is dimensionless. Now we add a second real scalar field, and we change the sign of the mass term, so that it is no longer a physical mass $m^2 = \omega^2 - \vec{k}^2$ for the particles.

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi_1 \partial^\mu \phi_1 + \partial_\mu \phi_2 \partial^\mu \phi_2 + \mu^2(\phi_1^2 + \phi_2^2) - \frac{\lambda}{2}(\phi_1^2 + \phi_2^2)^2\right). \tag{8.46}$$

This Lagrangian is invariant under any rotation of the fields of the form:

$$\begin{pmatrix} \phi_1' \\ \phi_2' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}. \tag{8.47}$$

The potential $V(\phi)$ is cylindrically symmetrical about the z axis in $x = \phi_1, y = \phi_2$ space:

$$V(\phi) = -\frac{\mu^2}{2}(\phi_1^2 + \phi_2^2) + \frac{\lambda}{4}(\phi_1^2 + \phi_2^2)^2. \tag{8.48}$$

Let $\phi_2 = 0$, then the maxima and minima of $V(\phi_1)$ are

$$\frac{\partial V}{\partial \phi_1} = (-\mu^2 + \lambda\phi_1^2)\phi_1 = 0; \quad \phi_1 = 0; \quad \phi_1 = \frac{\mu}{\sqrt{\lambda}}. \tag{8.49}$$

There are two points in the $\phi_1 - V(\phi)$ plane where the slope is zero: at the origin $V(0) = 0$, and at $\phi_1 = \mu/\sqrt{\lambda}$; $V(\phi) = -\mu^4/(4\lambda)$. For a particle inside the potential the origin is an unstable equilibrium, and the particle will roll off to the lower well, but can occupy any

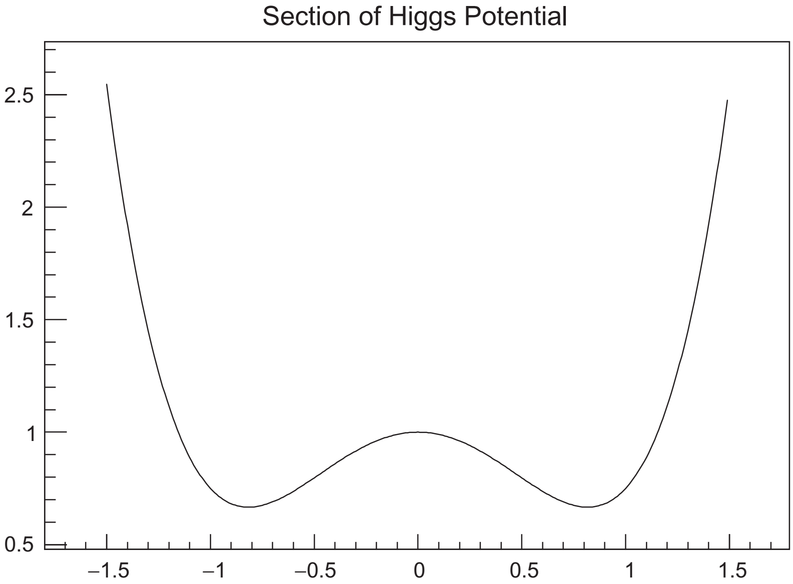


Figure 8.1: Sketch of a section of a Higgs potential. The function plotted is $V(\phi) = 1 - \phi^2 + 0.75\phi^4$. The potential is cylindrically symmetrical about the y axis. The vev is chosen in the well around $\phi = 0.85$ on the plot. Excitations in the radial direction bump into the well and acquire a mass. Excitation in the azimuthal direction meets no resistance, remains massless and becomes a Goldstone boson.

azimuth, because of the rotational symmetry. See Figure 8.1, $V(\phi)$ has the shape of a Mexican hat, with the crown at the origin, and the brim going to infinity. We may break the rotational symmetry by choosing a vacuum expectation value (vev) at a particular azimuth in the minimum of the hat, say:

$$\phi_2 = 0; \quad \phi_1 = \frac{\mu}{\sqrt{\lambda}} = v. \tag{8.50}$$

Then with respect to this origin, we define new fields:

$$\phi_1(x) \to v + \sigma(x); \quad \phi_2(x) \to \pi(x). \tag{8.51}$$

The vacuum expectation values of both fields now vanish, as is suitable in quantum field theory: $\langle 0|\,\sigma\,|0\rangle = \langle 0|\,\pi\,|0\rangle = 0$. Substituting these new fields into the Lagrangian and dropping constant terms gives:

$$\frac{1}{2}(\partial_\mu\sigma\partial^\mu\sigma + \partial_\mu\pi\partial^\mu\pi) + (\mu^2 v - \lambda v^3)\sigma + \frac{\mu^2}{2}(\sigma^2 + \pi^2) - \frac{3\lambda v^2}{2}\sigma^2$$
$$- \frac{\lambda}{4}(\sigma^4 + \pi^4) - \frac{\lambda\pi^2}{2}(v^2 + 2v\sigma + \sigma^2). \tag{8.52}$$

With $v = \mu/\sqrt{\lambda}$ the term linear in the σ field vanishes. Focus your attention on the term $(-3\lambda v^2/2)\sigma^2$. The minus sign comes from the potential, which is $-V(\phi)$ in the Lagrangian (remember $L = T - V$). The $-3\lambda v^2/2 = -3\mu^2/2$ combined with $+\mu^2/2$ flips the sign on the σ^2 term. Collecting terms gives the form:

$$\mathcal{L} = \frac{1}{2}\Bigg(\partial_\mu\sigma\partial^\mu\sigma + \partial_\mu\pi\partial^\mu\pi - 2\mu^2\sigma^2 - \frac{\lambda}{2}(\sigma^4 + \pi^4)$$
$$- 2\sqrt{\lambda}\mu(\sigma^3 + \pi^2\sigma) - \lambda\sigma^2\pi^2\Bigg). \tag{8.53}$$

The σ field has acquired a mass: $m_\sigma = \sqrt{2}\mu$. This field in our more complicated example will be the Higgs. The π field remains massless. The σ field with respect to the vev point oscillates radially, and hence bumps into the potential, giving rise to the mass. On the other hand the π field moves in the orthogonal direction, which freely goes around the hat in a circle, with no mass. The π field is called a 'Goldstone Boson', after J. Goldstone [4], who proved that whenever a symmetry is spontaneously broken, a massless boson is created. Expansion of the σ field about the symmetry breaking vev has generated a mass, and this is a called the 'Higgs mechanism'.

Our first example is a scalar field with a rotational symmetry that could be broken by choosing a vacuum expectation value away from the origin in ϕ space that generated a mass for one component of the field, leaving the other component massless. This is all very nice, but we need something more complicated to generate two masses for vector gauge fields, one for $W^\pm$ and one for Z.

Our second example increases the complexity in two ways. The symmetry is an abelian gauge transformation, and we add a new massless vector field, that we call B_μ, like the field in the neutral mix outlined above. Let $\rho = g'Y_W/2$ be the hypercharge of the field, and let the Higgs sector scalar bosons be hypercharged, so that there is a covariant derivative for ϕ that involves B. Recall from Chapter 3 that a hypercharged scalar field must be complex, like the $K^0\bar{K}^0, K_S K_L$. So, our Lagrangian is

$$\begin{aligned}
&\mathcal{L} = -\frac{1}{4}B^{\mu\nu}B_{\mu\nu} + (D_\mu\phi)^* D^\mu\phi - V(\phi), \quad D_\mu = \partial_\mu + i\rho B_\mu,\\
&B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad V(\phi) = -\mu^2\phi^*\phi + \frac{\lambda}{2}(\phi^*\phi)^2,\\
&\phi = \frac{\phi_1 + i\phi_2}{\sqrt{2}}.
\end{aligned} \tag{8.54}$$

The constants μ and λ conform to the definitions used by the Particle Data Group for the electroweak model. The covariant derivative is linear in the Lagrangian for the Dirac equation, but is quadratic for a scalar field that obeys the Klein–Gordon equation. The abelian gauge transformation is

$$\begin{aligned}
B_\mu(x) &\to B_\mu(x) + \partial_\mu\chi(x),\\
\phi(x) &\to e^{-i\rho\chi(x)}\phi(x),\\
D'_\mu\phi' &= e^{-i\rho\chi(x)}D_\mu\phi.
\end{aligned} \tag{8.55}$$

The scalar field ϕ has two symmetries, the gauge transformation, and rotation in (ϕ_1, ϕ_2) space identical to the example above. Following our previous analysis, we choose $\phi_2 = 0$, and $\phi_1 = \sqrt{2}\mu/\sqrt{\lambda}$, so that $v = \langle 0|\,\phi\,|0\rangle = \mu/\sqrt{\lambda}$. Substituting $\phi(x) = v + (\phi_1 + i\phi_2)/\sqrt{2}$ into $V(\phi)$ gives:

$$V(\phi) = -\frac{\mu^4}{2\lambda} + \mu^2\phi_1^2 + O(\phi^3) \text{ terms.} \tag{8.56}$$

So, the scalar field ϕ acquires a mass $m_\phi = 2\mu$, with ϕ_2 a massless Goldstone boson. Nothing new so far, but the B field also acquires a mass, as can be demonstrated through the covariant derivative.

$$\begin{aligned}
D_\mu\phi &= (\partial_\mu + i\rho B_\mu)\left(v + \frac{\phi_1 + i\phi_2}{\sqrt{2}}\right)\\
&= \frac{\partial_\mu(\phi_1 + i\phi_2)}{\sqrt{2}} + i\rho v B_\mu + i\rho B_\mu\frac{\phi_1 + i\phi_2}{\sqrt{2}},
\end{aligned}$$

and

$$(D^\mu\phi)^* = \frac{\partial^\mu(\phi_1 - i\phi_2)}{\sqrt{2}} - i\rho v B^\mu - i\rho B^\mu \frac{\phi_1 - i\phi_2}{\sqrt{2}}. \tag{8.57}$$

So, we have:

$$\begin{aligned} D_\mu\phi(D^\mu\phi)^* = {} & \frac{1}{2}D_\mu(\phi_1 + i\phi_2)(D^\mu(\phi_1 + i\phi_2))^* - iv\rho B^\mu D_\mu \frac{\phi_1 + i\phi_2}{\sqrt{2}} \\ & + \left(D^\mu \frac{\phi_1 + i\phi_2}{\sqrt{2}}\right)^* (i\rho v B_\mu) + \rho^2 v^2 B_\mu B^\mu. \end{aligned} \tag{8.58}$$

The last term is a mass term for the B_μ vector field. There is some cancellation in the second and third terms, leaving:

$$\begin{aligned} D_\mu\phi(D^\mu\phi)^* = {} & \frac{1}{2}D_\mu(\phi_1 + i\phi_2)(D^\mu(\phi_1 + i\phi_2))^* + \sqrt{2}\rho v B^\mu \partial_\mu \phi_2 \\ & + \sqrt{2}\rho^2 v B^\mu B_\mu \phi_1 + \rho^2 v^2 B_\mu B^\mu. \end{aligned} \tag{8.59}$$

The B_μ mass is

$$m_B^2 = \frac{2\rho^2\mu^2}{\lambda}, \tag{8.60}$$

where ρ is the hypercharge coupling constant, which enters the created vector boson mass term by way of the covariant derivative of the scalar field. It does not enter the mass formula for the scalar field itself. Note that only the hypercharge and the vev appear in the gauge boson mass formula, a feature that can be exploited to solve for the masses of the non-abelian gauge fields without the Higgs field (ϕ_1) in the equations. The arbitrary constants μ and λ may be determined by measuring the two masses m_ϕ and m_B, provided that the coupling constant ρ is known. However, one mass does not enable you to predict the other, a feature that made the search for the Higgs difficult: The next to the last term in Equation (8.59) gives the coupling of the ϕ field to B_μ.

The Goldstone boson ϕ_2 could be gauged away. The gauge transformation is

$$\phi' = e^{-i\rho\chi(x)}\phi = \frac{\cos(\rho\chi)\phi_1 + \sin(\rho\chi)\phi_2}{\sqrt{2}}, \quad \tan(\rho\chi) = \frac{\phi_2}{\phi_1}. \tag{8.61}$$

The tangent formula kills the imaginary part everywhere in space–time. On the other hand, the Goldstone boson in this example has a role to play as a needed degree of freedom. The massless B_μ field has no polarization component $\epsilon^{\lambda=0}$ in Equation (7.25), but the massive boson does, so the Goldstone boson could be 'eaten' by the B to become the third component of its polarization vector. It is this peculiar phenomenon that makes the large $\epsilon^{\lambda=0}$ term for W bosons in the decay of the top quark interesting.

After these two warmer-uppers, well really only one warmer-upper, since the second one includes the first, we are ready to take a look at the real thing. The gauge symmetry is non-abelian, and there is an independent abelian symmetry added on, the operators T_i and Y of our previous discussion. The scalar field must be hypercharged as well as a weak I-spin doublet, so it is a doublet of complex fields. For now, we leave the massless fermions out of the Lagrangian. As noted above, the gauge field masses come from the covariant derivative of the scalar. Choose $T_3 = -1/2$ and $Y = 1$ for the scalar vacuum. The gauge transformations on ϕ are:

$$\phi(x) \to e^{-ig\chi_i(x)T_i} e^{-g'\omega(x)/2}\phi(x), \tag{8.62}$$

with vacuum expectation value:

$$\langle 0|\,\phi\,|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}. \tag{8.63}$$

We may write the covariant derivative in terms of the fields W_i^μ and B^μ, because, as we shall soon see, the analysis has been carefully crafted by others to give us the answer we are looking for. Thus,

$$D_\mu = \partial_\mu + ig(W_{+\mu}T_+ + W_{-\mu}T_-) + igW_{3\mu}T_3 + ig'\frac{B_\mu}{2}. \tag{8.64}$$

In order to get the masses, we are interested in evaluating an expectation value of the square of this covariant derivative. The partial derivative terms do not contribute to the mass:

$$\begin{aligned}&\frac{1}{2}\begin{pmatrix}0 & v\end{pmatrix}\left\{ig(W_{+\mu}T_+ + W_{-\mu}T_-) + i\left(gW_{3\mu}T_3 + g'\frac{B_\mu}{2}\right)\right\}\\ &\quad\times\left\{ig(W_+^\mu T_+ + W_-^\mu T_-) + i\left(gW_3^\mu T_3 + g'\frac{B^\mu}{2}\right)\right\}\begin{pmatrix}0\\ v\end{pmatrix}.\end{aligned} \tag{8.65}$$

The operators give the following weak I-spin vectors:

$$T_+ \begin{pmatrix} 0 \\ v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} v \\ 0 \end{pmatrix}, \quad T_- \begin{pmatrix} 0 \\ v \end{pmatrix} = 0, \quad T_3 \begin{pmatrix} 0 \\ v \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 0 \\ v \end{pmatrix},$$
$$\begin{pmatrix} 0 & v \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix} = v^2. \tag{8.66}$$

Multiplying it all out gives:

$$\frac{v^2}{8}(-2g^2 W_{-\mu} W_+^\mu - (gW_{3\mu} - g' B_\mu)(gW_3^\mu - g' B^\mu)). \tag{8.67}$$

From Equations (8.38) and (8.40) we have:

$$gW_{3\mu} - g' B_\mu = \frac{gZ_\mu}{\cos\theta_W}, \text{ and } W_{-\mu} W_+^\mu = \frac{W_{+\mu}^* W_+^\mu + W_{-\mu}^* W_-^\mu}{2}. \tag{8.68}$$

Hence the vector boson masses are

$$M_W^2 = \frac{g^2 v^2}{4}, \quad M_Z^2 = \frac{g^2 v^2}{4\cos^2\theta_W},$$
$$M_W^2 = M_Z^2 \cos^2\theta_W. \tag{8.69}$$

The equation for M_W, plus our previous relation $g^2/(8M_W^2) = G_F/\sqrt{2}$ and some algebra gives:

$$v^2 = \frac{1}{\sqrt{2}G_F},$$
$$v = \frac{\mu}{\sqrt{\lambda}} = 246 \text{ GeV}. \tag{8.70}$$

The electromagnetic vector potential A_μ does not appear in our analysis because it is not included in the covariant derivative of the Higgs field. The scalar field ϕ has weak I-spin 1/2, and $Y = 1$, so the $T_3 = +1/2$ member is positively charged, and the $T_3 = -1/2$ member of $\phi^\dagger$ is negatively charged. These are Goldstone bosons that become the third component of the spin angular momentum of the massive $W^\pm$. The Higgs field is electrically neutral, and assigned $T_3 = -1/2$, and $Y = 1$. Therefore by design the electromagnetic vector potential does not occur in the Higgs covariant derivative, and does not acquire a mass by way of the Higgs mechanism. There is a third Goldstone boson that is neutral, the imaginary part of the Higgs field, that supplies the third component of the spin angular momentum for the Z. Equation (8.70) shows that the ratio $\mu/\sqrt{\lambda}$ is known from the Fermi coupling constant G_F, and that it has units of energy, the same as μ. With the discovery of the Higgs at $m_h = \sqrt{2}\mu = 125$ GeV, we now know the dimensionless parameter $\lambda = 0.13$.

8.5. Higgs Couplings

The Higgs field Lagrangian is

$$\mathcal{L} = (D_\mu\phi^\dagger D^\mu\phi) + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2, \tag{8.71}$$

where

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad \phi^\dagger(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & v + h(x) \end{pmatrix}. \tag{8.72}$$

Both the constant vev and the Higgs field are taken to be real. Therefore, we may calculate the Lagrangian:

$$\begin{aligned} \phi^\dagger\phi &= \frac{1}{2}(v^2 + 2vh + h^2), \\ (\phi^\dagger\phi)^2 &= \frac{1}{4}(v^4 + 4v^3h + 6v^2h^2 + 4vh^3 + h^4). \end{aligned} \tag{8.73}$$

Dropping the constant terms and noting that the term linear in the Higgs field cancels out, we have for the Lagrangian:

$$\mathcal{L} = (D_\mu\phi^\dagger D^\mu\phi) - \mu^2h^2 - \lambda vh^3 - \frac{\lambda h^4}{4}. \tag{8.74}$$

We have cubic and quardic self interaction terms. The coupling constant for the cubic term, describing the process $h \to h + h$ is $\lambda v = m_h^2/2v$. Combinatorics for the Feynman graph of $h \to h + h$ multiplies this coupling by $3! = 6$, so the Feynman rule has $3m_h^2/v$ at the vertex. The Higgs mass term is $\mu^2 = m_h^2/2$, or $m_h = \sqrt{2}\mu = 125$ GeV.

We return to the covariant derivative to obtain the couplings of the Higgs to the $W^{\pm}$ and Z vector bosons.

$$\begin{aligned} D_\mu\phi &= (\partial_\mu + ig(W_{+\mu}T_+ + W_{-\mu}T_-) \\ &\qquad + igW_{3\mu}T_3 + ig'B_\mu/2)\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h(x) \end{pmatrix} \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} i\frac{g}{\sqrt{2}}W_{+\mu}(v+h(x)) \\ \partial_\mu h(x) - i\frac{gZ_\mu}{2\cos\theta_W}(v+h(x)) \end{pmatrix}, \end{aligned} \tag{8.75}$$

where we have used $gW_{3\mu} - g'B_\mu = (gZ_\mu)/\cos\theta_W$. The expression in the Lagrangian $(D_\mu\phi)^\dagger D^\mu\phi$ is this column vector times its conjugate row vector:

$$\begin{aligned} (D_\mu\phi)^\dagger D^\mu\phi &= \frac{1}{2}\left(-i\frac{g}{\sqrt{2}}W_-^\mu(v+h(x))\ \partial^\mu h(x) + i\frac{gZ^\mu}{2\cos\theta_W}(v+h(x))\right) \\ &\qquad \times \begin{pmatrix} i\frac{g}{\sqrt{2}}W_{+\mu}(v+h(x)) \\ \partial_\mu h(x) - i\frac{gZ}{2\cos\theta_W}(v+h(x)) \end{pmatrix} \\ &= \frac{1}{2}\left\{\frac{g^2}{4}(W_{+\mu}^* W_+^\mu + W_{-\mu}^* W_-^\mu)(v+h(x))^2 \right. \\ &\qquad \left. + \partial_\mu h(x)\partial^\mu h(x) + \frac{g^2 Z_\mu Z^\mu}{4\cos^2\theta_W}(v+h(x))^2\right\}. \end{aligned} \tag{8.76}$$

The v^2 terms are the mass terms that we have obtained before, and the $vh(x)$ terms are the linear couplings of the Higgs to the vector bosons:

$$\begin{aligned} \mathcal{L}_{\text{int}} &= 2\frac{M_W^2}{v}W_\mu W^\mu h(x) + \frac{M_Z^2}{v}Z^\mu Z_\mu h(x) \\ &= (\sqrt{2}G_F)^{1/2}(2M_W^2 W_\mu W^\mu + M_Z^2 Z_\mu Z^\mu)h(x). \end{aligned} \tag{8.77}$$

The Higgs couplings are proportional to the squares of the masses. Higgs to $h + h \sim m_h^2$, $h \to Z + Z \sim M_Z^2$, and $h \to W + W \sim M_W^2$.

Up to now there have been no spin 1/2 Dirac particles in our electroweak Lagrangian. We have to insert what is called a Yukawa interaction term between the leptons and the Higgs. Following Barger and Phillips, we add a product interaction Lagrangian with a new unknown coupling constant:

$$\mathcal{L} = -G_e(\bar{u}_R \Phi^\dagger u_L + \bar{u}_L \Phi u_R),\ u_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L,$$
$$u_R = \text{right-handed electron}; \tag{8.78}$$

and

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{8.79}$$

Substitution of Φ in the Lagrangian gives:

$$\mathcal{L} = \frac{G_e}{\sqrt{2}}(\bar{e}_R e_L + \bar{e}_L e_R)(v + h(x)), \quad \text{giving } m_e = \frac{G_e v}{\sqrt{2}},$$

and the coupling of the Higgs to electrons is

$$\mathcal{L}_{\text{eff}} = (\sqrt{2}G_F)^{1/2} m_e \bar{e} e h. \tag{8.80}$$

Note that the Higgs coupling to the intermediate bosons is $h \to WW \sim M_W^2$, while the Higgs coupling to the electron, or any spin 1/2 particle is $h \to e^- e^+ \sim m_e$. This follows from the quadratic form of the Klein–Gordon equation and the linear form of the Dirac equation. From this effective Lagrangian the T matrix element for $h \to f\bar{f}$ (the Feynman rule for incoming or outgoing scalar particle is to multiply by one) is

$$\langle f\bar{f} | T | h \rangle = (\sqrt{2}G_F)^{1/2} m_f \bar{u}(p_1) v(p_2), \quad (p_1 + p_2)^2 = s = m_h^2. \tag{8.81}$$

In this notation, e_L represents a left-handed electron spinor $(1 - \gamma^5)/2\ u(p)$. Since the constant G_e is unknown, the model does not predict the electron mass. However, given the electron mass, the coupling to the Higgs m_e/v is known, and it is very small. The quarks are treated the same way, so only the top quark has much of a coupling constant. This means that the Drell–Yan process discussed extensively for W and Z production at $\bar{p}p$ and pp colliders, namely $\bar{u} + u \to h$, is very small for the Higgs. Drell–Yan production of an off-shell W that radiates a Higgs does make a contribution to Higgs production, but the dominant cross-section in $p\bar{p}$ or pp collisions is

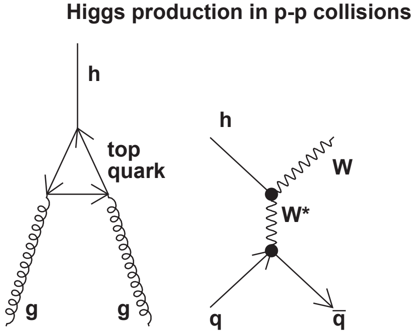


Figure 8.2: On the left two gluons fuse to create a top quark pair loop that annihilates to form a Higgs boson. The scalar Higgs is represented by a line. The right-hand figure is a Drell–Yan process that produces an off shell W^* that radiates a Higgs and becomes a real W. Gluon–gluon fusion is the dominant process for both the Tevatron and the LHC. At the Tevatron $W^* \to W + h$ is about 10% of the total, decreasing to 3% at 8 TeV.

called gluon–gluon fusion to form a $t\bar{t}$ pair that annihilate to make a Higgs. The Higgs does not couple directly to gluons. Both processes are diagrammed in Figure 8.2. Cross-section at the Tevatron for production of a 125 GeV Higgs is about 1.2 pb, while the $t\bar{t}$ cross-section is about 6 pb. At 13 TeV at the LHC the gluon–gluon fusion cross-section is about 55 pb, an increase of a factor of 50 [5]. The discovery channel for the Higgs was $h \to \gamma + \gamma$. Recent data from ATLAS at CERN are shown in Figure 8.3.

The Higgs mass 125 GeV is below threshold for $h \to t\bar{t}$ or $h \to W^+W^-$ or $h \to ZZ$. Both decays $h \to WW^* \to \nu\mu\nu\mu$ and $h \to ZZ^* \to 2\mu 2e$, where one vector boson is a propagator, have been seen, and will be discussed below. For Higgs decay the largest lepton branching ratio is $h \to \tau^+\tau^-$, and the largest accessible quark branching ratio is $h \to b\bar{b}$. We may calculate the decay rate for $h \to$ quarks from Equation (8.81).

$$\sum_{\text{spins}} |\langle q\bar{q}| T |h\rangle|^2 = \sqrt{2} G_F m_q^2 \mathrm{Tr}(\not{p}_1 + m_q)(\not{p}_2 - m_q)$$
$$= 4\sqrt{2} G_F m_q^2 (p_1 \cdot p_2 - m_q^2). \tag{8.82}$$

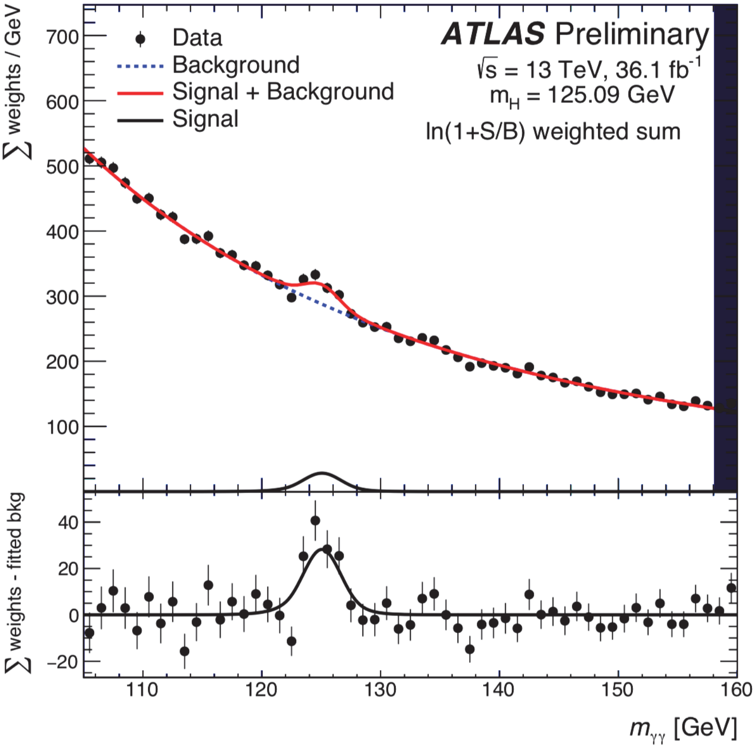


Figure 8.3: Atlas Collaboration $\gamma\gamma$ mass distribution for 36/fb of data taken at $\sqrt{s} = 13\,\text{TeV}$. The Higgs signal at 125 GeV is clearly seen in the background subtracted lower plot. Credit ATLAS-CONF-2017-045.

Using the kinematic relation, $p_1 \cdot p_2 - m_q^2 = (m_h^2/2)(1 - 4m_q^2/m_h^2)$ the squared matrix element becomes

$$\sum_{\text{spins}} |\langle q\bar{q}|\, T\, |h\rangle\, |^2 = 2\sqrt{2} G_F m_q^2 m_h^2 \left(1 - \frac{4m_q^2}{m_h^2}\right). \tag{8.83}$$

The formula for a two-body decay rate is

$$d\Gamma = \frac{|\langle q\bar{q}|\, T\, |h\rangle\, |^2}{2m_h (2\pi)^2} \delta^4(p - p_1 - p_2) \frac{d^3p_1}{2E_1} \frac{d^3p_2}{2E_2},$$

$$\frac{d\Gamma}{d\Omega} = \frac{G_F m_q^2 m_h}{16\pi^2\sqrt{2}} \left(1 - \frac{4m_q^2}{m_h^2}\right)^{3/2}. \tag{8.84}$$

Multiplying by 4π for the solid angle, three for color, gives the decay of the Higgs to quarks:

$$\begin{aligned}\Gamma_0(h \to q\bar{q}) &= \frac{3G_F m_h}{4\pi\sqrt{2}} m_q^2 \left(1 - \frac{4m_q^2}{m_h^2}\right)^{3/2} \\ &= 4.3 \text{ MeV for } m_b = 4.2 \text{ GeV}. \end{aligned} \tag{8.85}$$

Even for the b quark $4m_q^2/m_h^2 = 0.0045$, so that factor can be set to one without penalty. The question is what to use for the quark masses? Almost all of the formulas derived in this book are accurate to a few %, but in this case if one uses the PDG quark masses $m_b = 4.2$ GeV, $m_c = 1.3$ GeV, and $m_s = 0.1$ GeV one is off a factor of 2! The Higgs has no strong interactions, but the quarks do, and the quark energy is $m_h/2 = 62$ GeV, a factor of 10 higher than m_b. This means that there is a substantial correction to the effective quark mass from QCD final state interactions. This effect, that makes the quark effective mass smaller, was first calculated back in 1980 [6]. If Γ_0, Equation (8.85), is the uncorrected rate, then the corrected rate Γ is given by

$$\Gamma = \Gamma_0(\ln(2m_q/\Lambda)/\ln(m_h/\Lambda))^{24/(33-2n_f)} \left(1 + 5.67\frac{\alpha_s(m_h^2)}{\pi}\right). \tag{8.86}$$

The first correction term is quark mass dependent, so it is different for different flavors. The second term, involving $\alpha_s(m_h^2)$, is common to all quark flavors. Equation (5.95) in the discussion of the parton distribution functions, defined the running coupling constant α_s for QCD as

$$\alpha_s(Q^2) = \frac{2\pi}{(11 - 2n_f/3)\ln(Q/\Lambda)}. \tag{8.87}$$

In these equations, $n_f = 5$ is the number of quark flavors (u, d, s, c, b), and Λ is the cut-off energy where $\alpha_s \to \infty$. The mass dependent part of Equation (8.86) was obtained from diagrams proportional to α_s by a procedure called the leading logarithm approximation. Braaten and Levielle used $\Lambda = 0.5$ GeV. The exponent in Equation (8.86) can safely be set to one (24/23). The 'running mass' at the energy of the Higgs is smaller than the mass that is obtained from the $\Upsilon(b\bar{b})$ bound state (see Chapter 9). The common correction term is about 1.25. Then the corrected rates are

$$\Gamma_b = 0.51 \times 1.25\Gamma_{0b} = 2.7\,\text{MeV}, \quad \Gamma_c = 0.30 \times 1.25\Gamma_{0c} = 0.15\,\text{MeV}. \tag{8.88}$$

The Γ_0 formula should be good for the charged leptons (Higgs does not couple to massless neutrinos), without the color factor of three and $m_\tau = 1.8$ GeV, one gets:

$$\begin{aligned} \Gamma(h \to \tau^+\tau^-) &= 0.26\ \text{MeV}, \\ \Gamma(b\bar{b})/\Gamma(\tau\tau) &= 10.4, \\ \Gamma(\tau\tau)/\Gamma(\mu\mu) &= \frac{m_\tau^2}{m_\mu^2} = 283. \end{aligned} \tag{8.89}$$

The vector boson final states contribute about 2 MeV, so the total width of the Higgs is about: $\Gamma(h \to all) \sim 5$ MeV. This is a very narrow width for a particle with 125 GeV mass. It is about 1/500 of the W and Z widths, which have masses in the same range, and share the same weak interaction.

The Higgs was discovered at the LHC via the two body decay $h \to \gamma + \gamma$ that has a branching ratio of about 0.2%, so it is a rare decay, but with good mass resolution. See the plots in Figure 8.3, where there are about 120 events in the signal peak. The two-body modes $h \to b\bar{b}$ and $h \to \tau^+\tau^-$ have both also been seen, but the mass resolution is degraded by the hadronization of the b quarks, and the fact that the decay of the τ lepton always involves a neutrino in the final state. The decay $h \to \mu^-\mu^+$ should be about 10% of the discovery channel $h \to \gamma\gamma$. The dimuon channel has better mass resolution than $\gamma\gamma$. The CMS collaboration at the CERN LHC reported a dimuon signal at the Higgs mass in 2020, but the statistical precision is limited, and more data are needed to improve on the Higgs mass. (cms.cern/news/cms-sees-evidence-higgs-boson-decaying-muons). The theoretical expectation for the Higgs branching ratios, taken from the Particle Data Group, are listed in Table 8.3, and compared to our first-order calculations.

Table 8.3: Theoretical predictions for Higgs decay branching ratios.

Higgs decay channel	Branching ratio	Rate (MeV)
$\gamma\gamma$	2.3×10^{-3}	
ZZ^*	2.6×10^{-2}	
WW^*	2.1×10^{-1}	1.44 (Eq. (8.102))
$\tau\tau$	6.3×10^{-2}	0.26 (Eq. (8.84))
$b\bar{b}$	5.8×10^{-1}	2.7 (Eq. (8.84))*
$c\bar{c}$	2.9×10^{-2}	0.15 (Eq. (8.84))*
$Z\gamma$	1.5×10^{-3}	
$\mu\mu$	2.1×10^{-4}	5×10^{-4} (Eq. (8.84))

Asterisk signifies quark mass correction via Equation (8.86).

Agreement between the PDG numbers and ours is good for the fermions. The ratio $b\bar{b}/c\bar{c} = 18$ from our formulas, and is 20 for PDG. Similarly we get $b\bar{b}/\tau\tau = 10.4$, while the PDG number is 9.2. One might expect the purely leptonic rates to be trouble free, and in fact in the table $\tau\tau/\mu\mu = 300$, while $m_\tau^2/m_\mu^2 = 283$. All of these results are pretty close. The absolute decay rates into fermions are also in reasonable agreement. For instance, for $h \to b\bar{b}$ $2.7\,\text{MeV}/0.58 = 4.6\,\text{MeV}$ full width. The rate for $h \to WW^* =$ 1.44 MeV calculated below (Equations (8.94)ff) seems to be a bit large relative to the PDG branching ratio, since $1.44/0.21 = 6.8\,\text{MeV}$ total width.

As we pointed out in Chapter 7 in the discussion of $p\bar{p} \to W, Z$, the experimental measurement is the cross-section times the branching ratio for a particular decay mode, so that the individual decay rates are not directly measured. Decays to different final states will give ratios, because the production cross-section cancels out. The same problem obtains for the Higgs. The LHC experiments can directly measure a branching ratio by assuming that the theoretical prediction for the production cross-section is correct. The total width of the Higgs is so tiny that there is no chance of obtaining it from the width of a mass peak that is dominated by experimental resolution. The full width of the $\gamma\gamma$ peak in Figure 8.3 is about 4 GeV. So, $\Gamma(h \to all)$ will have to come from a sum over theoretical calculations of the main partial decays, unless there is an observable proportional to Γ_{tot} that can be measured.

The narrow width of the Higgs is due to the fact that the mass is below threshold to decay to pairs of vector bosons W^+W^- and ZZ, so that for a final state $h \to \nu\bar{\nu}\mu^+\mu^-$ one of the W's must be a propagator, and that substantially affects the decay rate. Before we calculate the decay rate for a 125 GeV Higgs, let us suppose that the Higgs mass is 180 GeV, so that $h \to W^+W^-$ is energetically allowed (but ZZ is right on the edge). The rate will be several orders of magnitude larger than for the true Higgs mass. We return to the Lagrangian Equation (8.77):

$$\mathcal{L}_{\text{int}} = (\sqrt{2}G_F)^{1/2}(2M_W^2 W^\mu W_\mu)h(x),$$
$$\langle W^+W^- | \, T \, |h\rangle = (\sqrt{2}G_F)^{1/2} 2M_W^2 \epsilon_\mu^*(p_1,\lambda_1)\epsilon^{*\mu}(p_2,\lambda_2). \quad (8.90)$$

The two ϵ's are for the two final state vector bosons. The matrix element is simple, and involves no traces, since the various decay modes of the $W^\pm$ come automatically.

$$\sum_{\text{spins}} |\langle W^+W^- | \, T \, |h\rangle|^2$$
$$= 4M_W^4\sqrt{2}G_F \sum_{\lambda_1}\sum_{\lambda_2} \epsilon_\mu^*(p_1,\lambda_1)\epsilon_\nu(p_1,\lambda_1)\epsilon^{*\mu}(p_2,\lambda_2)\epsilon^\nu(p_2,\lambda_2). \quad (8.91)$$

Evaluating the spin sums $\sum \epsilon_\mu^*\epsilon_\nu = -g_{\mu\nu} + p_\mu p_\nu/M_W^2$ gives for $m_h =$ 180 GeV:

$$\sum_{\text{spins}} |\langle W^+W^- | \, T \, |h\rangle|^2 = \sqrt{2}G_F m_h^4(12\lambda_W^2 - 4\lambda_W + 1),$$
$$\lambda_W = \frac{M_W^2}{m_h^2} = 0.2. \quad (8.92)$$

Here, we have used $g^{\mu\nu}g_{\mu\nu} = 4$, $p_1^2 = p_2^2 = M_W^2$, and $(p_1 \cdot p_2)^2/M_W^4 = 1/(4\lambda_W^4) - 1/\lambda_W + 1$. Note that the multiplier M_W^4 has been replaced by m_h^4. Using the two-body decay rate formula Equation (8.84) gives

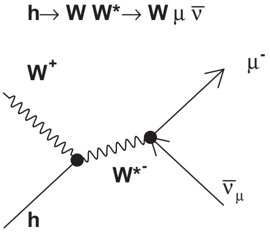


Figure 8.4: Diagram for a 125 GeV Higgs decaying into a real W^+ and a propagator W^{*-}. The W^* then decays into $\mu^-\bar{\nu}_\mu$. This reaction has a much smaller decay rate than $h \to WW$ for a Higgs mass greater than $2M_W$.

the answer:

$$\begin{aligned}\Gamma(h \to W^+W^-) &= \frac{1}{8\pi}\frac{G_F}{\sqrt{2}}m_h^3(1-4\lambda_W)^{1/2}(12\lambda_W^2 - 4\lambda_W + 1)\\ &= 585 \text{ MeV for 180 GeV Higgs.}\end{aligned} \tag{8.93}$$

Therefore, if the Higgs mass were above threshold for W^+W^- and ZZ decays, the full width would be in the 1 GeV range. But the Higgs mass is 125 GeV, so we have to solve a more difficult problem that involves the W propagator. The decay diagram, shown in Figure 8.4, has a three-body final state ($W\mu\nu$), and resembles muon decay

The transition matrix element from the diagram is

$$\begin{aligned}-i\langle W\mu\nu|\, T\, |h\rangle &= 2M_W^2(\sqrt{2}G_F)^{1/2}\frac{g}{\sqrt{2}}\epsilon_\mu^{*\lambda}\\ &\quad\times i\frac{-g^{\mu\nu}+q^\mu q^\nu/M_W^2}{q^2-M_W^2}\bar{u}(k)\gamma_\nu\frac{1-\gamma^5}{2}v(k').\end{aligned} \tag{8.94}$$

The four momenta are $p = p' + k + k'$, with p the Higgs, p' the W, and $k + k'$ the massless $\mu^-\bar{\nu}$ pair. The four momentum of the propagator is $q = p - p' = k + k'$. Evaluation of the decay rate from Equation (8.94) is straightforward, but a bit lengthy. The amplitude Equation (8.94) is proportional to G_F ($g \sim \sqrt{G_F}$), while Equation (8.90) for real $W^\pm$ is proportional to $\sqrt{G_F}$, and that suppresses the rate for the propagator case. The calculation is simplified by multiplying the Dirac current by the propagator in the amplitude, and using the Dirac equation to eliminate the $q^\mu q^\nu/M_W^2$ term. The traces involving γ^5 vanish. The spin averaged squared matrix element is

$$\sum_{\text{spins}} |\langle W\mu\nu| T |h\rangle|^2 = 32\frac{M_W^4}{m_h^4}\frac{M_W^2 G_F^2}{(1-2E'/m_h)^2} \times \left(k\cdot k' + 2\frac{k\cdot p' k'\cdot p'}{M_W^2}\right). \tag{8.95}$$

Substitution in the differential decay rate formula gives:

$$d\Gamma = \frac{1}{2\pi^5}\frac{M_W^6}{m_h^6}\frac{m_h G_F^2}{(1-2E'/m_h)^2}\left(k\cdot k' + 2\frac{(k\cdot p' k'\cdot p')}{M_W^2}\right) \times \delta^4(p-p'-k-k')\frac{d^3p'}{2E'}\frac{d^3k}{2\omega}\frac{d^3k'}{2\omega'}. \tag{8.96}$$

The matrix element depends on the four vectors k and k', so it is convenient to write the δ function as $\delta^4((p-p')-k-k')$, and use Equation (6.53) for two neutrinos, since we may neglect the muon mass:

$$d_2(PS) = \delta^4(X-k-k')\frac{d^3k}{2\omega}\frac{d^3k'}{2\omega'},$$

$$\int d_2(PS)k^\alpha k'^\beta = \frac{\pi}{24}(X^2 g^{\alpha\beta} + 2X^\alpha X^\beta), \tag{8.97}$$

here $X = p - p'$. The proof of this formula is outlined in Problem 6.7. Evaluating the two-body phase space gives the following formula in terms of the W four momentum:

$$d\Gamma = \frac{1}{4\pi^4}\frac{M_W^6}{m_h^6}\frac{m_h^3 G_F^2}{(1-2E'/m_h)^2}\frac{d^3p'}{2E'} \times (2/3 + (M_W/m_h)^2 - 2E'/m_h + E'^2/(3M_W^2)). \tag{8.98}$$

It is convenient to introduce the dimensionless variables $x = 2E'/m_h$ and $\epsilon = M_W/m_h$. Then after integration over the W solid angle for a factor of 4π, we have the formula as first derived by Keung and

Marciano [7]:

$$\frac{d\Gamma(h \to W\mu\nu)}{dx} = \frac{m_h^5 G_F^2}{96\pi^3} \frac{\epsilon^4 \sqrt{x^2 - 4\epsilon^2}}{(1-x)^2} (x^2 + 12\epsilon^4 - 12\epsilon^2 x + 8\epsilon^2). \tag{8.99}$$

Integration is a bit of a chore. My integral table does not have this form, but it does if you substitute $y = 1 - x$, that is to say that the integrals themselves are standard. It is the number of terms that makes it tedious. The limits are $x_1 = 2\epsilon$, and $x_2 = 1 + \epsilon^2$. Keung and Marciano write the final result as

$$\Gamma(h \to W\mu\nu) = \frac{m_h^5 G_F^2}{96\pi^3} \epsilon^4 F(\epsilon), \tag{8.100}$$

where

$$\begin{aligned} F(\epsilon) = {} & \frac{3(1 - 8\epsilon^2 + 20\epsilon^4)}{\sqrt{4\epsilon^2 - 1}} \cos^{-1}\left(\frac{3\epsilon^2 - 1}{2\epsilon^3}\right) \\ & - (1 - \epsilon^2)\left(\frac{47\epsilon^2}{2} - \frac{13}{2} + \frac{1}{\epsilon^2}\right) \\ & - 3(1 - 6\epsilon^2 + 4\epsilon^4)\ln(\epsilon). \end{aligned} \tag{8.101}$$

The factor $m_h^5 G_F^2$ is reminiscent of muon decay $m_\mu^5 G_F^2$. With $m_h =$ 125 GeV this term has no exponents: $m_h^5 G_F^2 = 4.1$ GeV. Doing the arithmetic gives the transition rate in MeV:

$$\begin{aligned} \Gamma(h \to W\mu\nu) &= 0.08 \text{ MeV}, \\ \Gamma(h \to WW^*) &= 1.44 \text{ MeV}. \end{aligned} \tag{8.102}$$

The difference between the two rates is a factor of 18 that takes into account all of the W^* decay modes, and the two signs of the charge. The ZZ^* cross-section is smaller by about 1/8, because of the coupling and the energy available.

Figure 8.4 has other applications. One proposal to study the Higgs is to build a new e^-e^+ collider with $\sqrt{s} = 250$ GeV or more, to be well above threshold for Higgs production via $e^-e^+ \to Z + h$. Figure 8.4 can describe this process. Imagine e^-e^+ incident, replacing the final state leptons $\mu\bar{\nu}$, and the propagator is a Z^* that decays into $Z + h$. The diagram simply goes in the opposite direction, with

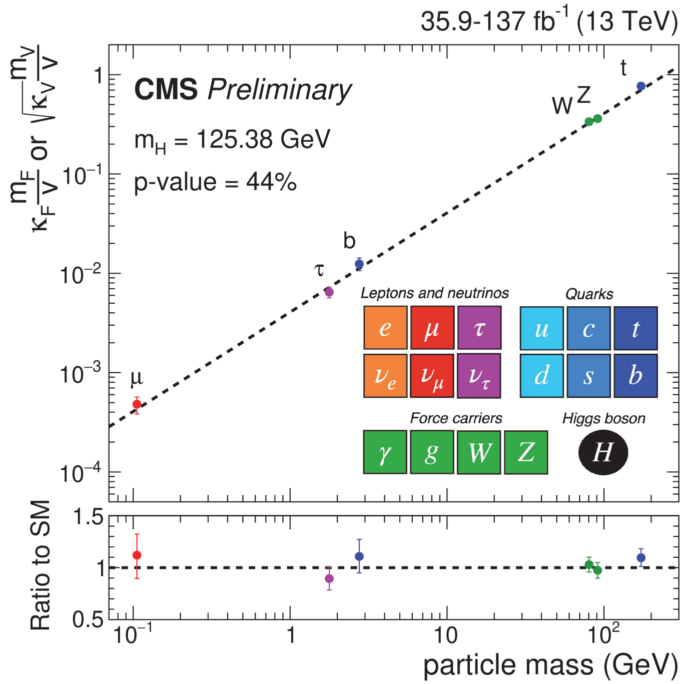


Figure 8.5: Measurements of the Higgs decay rates by the CMS Collaboration at the CERN Large Hadron Collider. The top plot shows the Higgs decay particle mass in units of $v = 246\,\text{GeV}$. The k factors are small corrections. The decay $h \to t\bar{t}$ is energy forbidden, but the coupling can be studied from the $ht\bar{t}$ final state. The bottom plot shows a comparison between the theoretical expectations of the electroweak theory with experiment. Credit CMS Collaboration, CERN.

W coupling replaced by Z coupling. The neutral e^-e^+ pair cannot produce a single W^*, but a pair of quarks can. So, Figure 8.4 can be used to describe the Higgs production process via Drell–Yan: $u\bar{d} \to W^{*+} \to W^+ + h$. This is one of the smaller contributors to Higgs production at the LHC (see the PDG article).

Experimental work on Higgs production and decay is still in progress in 2020. A snapshot of the situation has been furnished by the CMS Collaboration at CERN, and is shown in Figure 8.5.

Our treatment of the Higgs boson, a brand new elementary particle in the electroweak theory, is only first order, and as we have seen there are large corrections to some of the predictions. There is still a lot to learn about the Higgs, and the subject is very actively pursued. Even though the cross-section for Higgs production at 13 TeV is about 500 fb (5×10^{-37} cm^2), it is fair to say that the experiments are difficult. Even the top quark is a troublesome background for $h \to b\bar{b}$, because a $t\bar{t}$ event, with a larger cross-section than the Higgs, also has $b\bar{b}$ in the final state. Someone once observed that last year's discovery is this year's background. The reader is referred to the CERN site mentioned earlier, or the annual review by the Particle Data Group, for information regarding experimental and theoretical progress. The existence of the Higgs boson and its couplings to the masses of the other particles is an impressive confirmation of the electroweak theory.

8.6. *W* Magnetic Moment

In Chapter 2, in the discussion of the Dirac equation and the derivation of the magnetic moment of the electron the g factor was defined by the equation:

$$\vec{\mu} = \frac{ge}{2m}\vec{S}. \tag{8.103}$$

This is not to be confused with the coupling constant g introduced in the electroweak theory. We pointed out that it has been argued by some that $g = 2$, the result for Dirac's electron, is the 'natural' value for all charged elementary particles with spin. Besides the charged spin 1/2 leptons, there are only two other charged elementary particles that have no strong interactions, the $W^{\pm}$ vector bosons. We will now demonstrate that, because of the non-abelian gauge theory, the g factors for $W^{\pm}$ are also two.

In the preceding sections on the electroweak theory, the Higgs mechanism, the vector boson masses, the Higgs boson, and all of the couplings, the vector potential A_μ of QED was on purpose left out of our discussion. The Higgs mechanism was designed to leave A_μ massless. However, if we want to calculate the W magnetic moment, we have to resurrect the QED vector potential. It is buried in the

Lagrangian of Equations (8.34) and (8.35). The term of interest is

$$\mathcal{L} = -\frac{1}{4} W_i^{\mu\nu} W_{i\mu\nu}, \quad W_i^{\mu\nu} = \partial^\mu W_i^\nu - \partial^\nu W_i^\mu - g\epsilon_{ijk} W_j^\mu W_k^\nu. \tag{8.104}$$

This is all we will need, because W_3^μ and A^μ are related by the angle θ_W. Expanding the product gives:

$$\begin{aligned}\mathcal{L} = -\frac{1}{4}((\partial^\mu W_i^\nu - \partial^\nu W_i^\mu)^2 - 2g\epsilon_{ijk}(\partial_\mu W_{i\nu} - \partial_\nu W_{i\mu}) W_j^\mu W_k^\nu \\ + g^2((\vec{W}^\mu \cdot \vec{W}_\mu)(\vec{W}^\nu \cdot \vec{W}_\nu) - (\vec{W}^\mu \cdot \vec{W}_\nu)(\vec{W}^\nu \cdot \vec{W}_\mu)).\end{aligned} \tag{8.105}$$

The second term contains the interaction of the W spin with the magnetic field, as we shall see after some manipulation. The Dirac equation has the advantage that through the γ matrices the electron spin is built into the Lagrangian. The W spin is in its Lagrangian too, it is just not as obvious. The components of the vector field W^μ are proportional to the components of the polarization. So, we isolate and expand the second term:

$$\begin{aligned}&\frac{1}{2} g\epsilon_{ijk}(\partial_\mu W_{i\nu} - \partial_\nu W_{i\mu}) W_j^\mu W_k^\nu \\ &\quad = \frac{g}{2}((\partial_\mu W_{1\nu} - \partial_\nu W_{1\mu})(W_2^\mu W_3^\nu - W_3^\mu W_2^\nu) \\ &\qquad + (\partial_\mu W_{2\nu} - \partial_\nu W_{2\mu})(W_3^\mu W_1^\nu - W_1^\mu W_3^\nu) \\ &\qquad + (\partial_\mu W_{3\nu} - \partial_\nu W_{3\mu})(W_1^\mu W_2^\nu - W_2^\mu W_1^\nu)).\end{aligned} \tag{8.106}$$

The space components of the last term, after the identification $W_3 = \sin\theta_W A + \cos\theta_W Z$, and dropping the Z part, have the form $\vec{S} \cdot \vec{B}$, with $g = 1$. After integration by parts an identical second term can be extracted from the rest of the formula, giving $g = 2$ as a final answer.

Integration by parts can always be done on a Lagrangian density, because the constant terms vanish. Thus,

$$(\partial_\mu W_{1\nu}) W_2^\mu W_3^\nu = -W_{1\nu} W_3^\nu \partial_\mu W_2^\mu - W_{1\nu} W_2^\mu \partial_\mu W_3^\nu. \tag{8.107}$$

There are two terms $-W_{1\nu}W_2^\mu\partial_\mu W_3^\nu$ and $-W_{2\nu}W_1^\mu\partial_\mu W_3^\nu$. There are also two terms involving $\partial_\nu W_3^\nu$ which can be gauged away. The remaining terms involve derivatives of other components of the W field, and are not of interest for the magnetic moment. The useful part of Equation (8.107) after substitution for A is

$$eF_{\mu\nu}(W_1^\mu W_2^\nu - W_1^\nu W_2^\mu) = -2ieF_{\mu\nu}W_+^\mu W_-^\nu, \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{8.108}$$

where we have used $W_1 = (W_+ + W_-)/\sqrt{2}$, and $W_2 = i(W_+ - W_-)/\sqrt{2}$, and $F_{\mu\nu} = -F_{\nu\mu}$. The space components of Equation (8.108) are the answer to our problem.

$$\begin{aligned} -2ieF_{ij}W_{+i}W_{-j} &= -2ie\epsilon_{ijk}B_kW_{+i}W_{-j} \\ &= -2ie\vec{W}_+ \times \vec{W}_- \cdot \vec{B} = 2eB_z, \end{aligned} \tag{8.109}$$

where we have chosen $W_{+j} = \epsilon_j^{\lambda=+1}$, $W_{-j} = W_{+j}^*$, and $\epsilon^* \times \epsilon = \hat{z}$. This is equivalent to $\langle S_z \rangle = +1$. Dividing by $2M_W$ for the normalization of the W field gives the energy:

$$U = -\frac{e}{M_W}\vec{S}\cdot\vec{B}, \quad g = 2. \tag{8.110}$$

The minus sign comes from conversion to a potential energy from the Lagrangian. Remember $L = T - V$. If the extra terms from the non-abelian gauge invariance were absent, the answer would be $g = 1$, and one might expect that value for an ‘ordinary’ vector meson. However, there are no other known charged vector particles that have no strong interaction, and hence could be expected to have a g factor from first-order theory. For higher spin, the Ω^- has spin 3/2, but its magnetic moment is composite, made of three strange quarks. Could the $W^\pm$ g factor be measured? The W is very short lived ($\sim 3 \times 10^{-25}$ s), so it will not precess much in an external magnetic field, even a very strong one, before decaying. The magnetic moment depends on the triple gauge boson coupling of the form WWV, where $V = A$, the QED vector potential, and anomalous triple gauge couplings can be studied. The Particle Data Group has a nice review of the present experimental situation. They write the magnetic moment as $\mu_W = e(1 + k_\gamma + \lambda_\gamma)/(2M_W)$. In the standard model, which is the present analysis, the anomalous couplings are $k_\gamma = 1$ and $\lambda_\gamma = 0$, and experimental data are consistent with these numbers. Hence yes, indirectly the W magnetic moment can be measured, and is consistent with expectations of the standard model.

The author is indebted to his colleagues Barger and Phillips and Peskin and Schroeder, whose published work was followed in much of this chapter.

8.7. Problems

8.1 Derive the commutation rule for two covariant derivatives in QED, Equation (3.207):

$$[D_\mu, D_\nu] = ieF_{\mu\nu}. \tag{8.111}$$

8.2 Repeat the calculation for the Yang–Mills covariant derivative to obtain Equation (8.16).

8.3 Derive Equation (8.18), proving that the Dirac equation is gauge invariant in the Yang–Mills theory.

8.4 Prove that the free field Lagrangian Equation (8.14) is gauge invariant by deriving Equation (8.21) and demonstrating that Equation (8.22) proves invariance.

8.5 Obtain the Yang–Mills wave Equation (8.25) by taking the necessary functional derivatives of the Lagrangian: Equation (8.16) in Equation (8.26).

8.6 Obtain Equation (8.44) from Equation (8.42), demonstrating that the Z couplings used in Chapter 7 are correct.

8.7 Do the algebra to show that the Lagrangian Equation (8.53) for the σ, π model is correct.

8.8 Evaluate the expectation value of the square of the covariant derivative Equation (8.65) and obtain the vector boson masses Equation (8.69).

8.9 Calculate $(D_\mu\phi)^\dagger D^\mu\phi$ and obtain the couplings of the $W^\pm$ and Z to the Higgs, Equation (8.77).

8.10 Begin with Equation (8.81) and calculate the decay rate in MeV for $h \to b\bar{b}$. Use the PDG value for $m_b = 4.2\,\text{GeV}$.

8.11 Calculate the cross-section in the e^-e^+ center of mass for $e^-e^+ \rightarrow Z + h$ at $\sqrt{s} = 250$ GeV. Use the diagram of Figure 8.4 turned around, with e^-e^+ annihilation in the initial state to create a Z propagator that decays into $Z + h$. It is a two-body final state, which is simpler than $h \rightarrow WW^* \rightarrow W\mu\nu$. Use the appropriate coupling of the Z to electrons. Express your final answer in units of $R = (4\pi\alpha^2)/(3s)$. The Higgs cross-section is very small. At 250 GeV $R = 1.36$ pb, giving about 5 muon pairs per hour at 10^{33} luminosity. It is interesting to note that the same initial and final states could go via a Higgs propagator, which is $1/(p^2 - m_h^2 + i\epsilon)$: $e^-e^+ \rightarrow h^* \rightarrow h + Z$, except for the fact that the last step does not exist. The Z can radiate a Higgs, but the Higgs cannot radiate a Z and conserve angular momentum.

The amplitude is

$$\langle Zh| T \left|e^-e^+\right\rangle = \frac{g}{2\cos\theta_W} \frac{(\sqrt{2}G_F)^{1/2} M_Z^2 \epsilon_\nu^{*\lambda}}{s - M_Z^2}(-g^{\mu\nu}) \times \bar{v}(k')\gamma_\mu(g_L(1-\gamma^5) + g_R(1+\gamma^5))u(k). \tag{8.112}$$

A two-body final state in the center of mass has no free parameters other than the angles. The two momenta are equal and opposite, and the two energies are given by familiar formulas:

$$E_Z = \frac{s + M_Z^2 - m_h^2}{2\sqrt{s}}, \quad E_h = \frac{s + m_h^2 - M_Z^2}{2\sqrt{s}}. \tag{8.113}$$

The spin averaged square of the amplitude is

$$\sum_{\text{spins}} |\langle Zh| T \left|e^-e^+\right\rangle|^2 = \frac{16G_F^2 M_Z^6}{(s - M_Z^2)^2}(g_L^2 + g_R^2)\left(k \cdot k' + \frac{2}{M_Z^2}(k \cdot p')(k' \cdot p')\right), \tag{8.114}$$

where p' is the Z four momentum in the final state, and k and k' are the initial state four momenta for e^- and e^+.

$s = (k+k')^2 = (p+p')^2$. Integrate the differential cross-section over the Z solid angle to obtain:

$$\sigma = \frac{2}{3\pi}\frac{G_F^2 M_Z^4 |\vec{p}|}{\sqrt{s}}(g_L^2 + g_R^2)\frac{E_Z^2 + 2M_Z^2}{(s - M_Z^2)^2}, \tag{8.115}$$

where $|\vec{p}|$ is the momentum of either particle in the final state. Obtain the coupling constants for eeZ from Table 7.2, and numerically evaluate the cross-section in units of R. I got a cross-section $\sigma \sim 0.25$ pb, less than one unit of R. This puts pressure on the storage ring or linear collider to achieve the highest possible luminosity.

8.12 W magnetic moment. Do the parts integration on Equation (8.106) to obtain the two identical contributions to Equation (8.109). Then check that $\vec{\epsilon}^{\,\lambda=1\,*} \times \vec{\epsilon}^{\,\lambda=1} = \hat{z}$.

References

[1] C. N. Yang and R. Mills, *Phys. Rev.* **96**, 191 (1954).

[2] A. Erwin *et al.*, *Phys. Rev. Lett.* **6**, 628 (1961).

[3] S. Coleman, *Quantum Field Theory Lectures* (World Scientific, Singapore, 2019).

[4] J. Goldstone, *Nuovo Cimento* **19**, 154 (1961).

[5] M. Carena *et al.*, *Status of Higgs Boson Physics*, in P. A. Zyla *et al.*, (Particle Data Group), *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020). See also LHC Higgs Cross-section Working Group CERN report 2017-002 (2016), arXiv:1610.07922.

[6] E. Braaten and J. Levielle, *Phys. Rev. D* **22**, 715 (1980).

[7] W.-Y. Kueng and W. J. Marciano, *Phys. Rev. D* **30**, 248 (1984).

Chapter 9

Heavy Quark Bound States, Mixing, and CP Violation

9.1. Hydrogen Atom

A good place to begin a discussion of exotic atoms is to review one that is definitely not exotic, namely hydrogen. So that our formulas look like those in standard texts on quantum mechanics, we will return to cgs units, with the elementary charge $e = 4.8 \times 10^{-10}$ esu, and $\hbar c = 197\,\text{eV-nm}$. The Bohr radius is $a_0 = \hbar^2/(me^2) = 0.0529\,\text{nm}$, where m is the reduced mass $m = (m_e m_p)/(m_e + m_p)$. The central force two-body problem can be simplified to the relative motion of a reduced mass particle in the center of mass of the system. The radius r is the distance between the two bodies, which for hydrogen is very close to the distance from the electron to the center of mass because the proton is 2000 times heavier than the electron. The radial Schrödinger equation for the bound states of hydrogen is

$$\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr}R_{nl}(r) - \frac{2m}{\hbar^2}\left(-\frac{Ze^2}{r} + \frac{\hbar^2 l(l+1)}{2mr^2} - E_n\right)R_{nl}(r) = 0. \tag{9.1}$$

Although the radial wave function is different for different values of l, the energy eigenvalue is l degenerate:

$$\begin{aligned} E_n &= -\frac{Z^2\alpha^2 mc^2}{2n^2} = E_0 Z^2; \\ E_0 &= -\frac{13.6}{n^2}\ \text{eV}, \quad \text{Bohr's formula}; \\ n &= 1,2,3,\ldots;\ l = 0,1,2,\ldots,n-1. \end{aligned} \tag{9.2}$$

For each value of l there are $2l+1$ values of m_l, giving the total number of states with a given energy $N = \sum_0^{n-1}(2l+1) = n^2$ (remember the simple rule $\sum_1^n(n) = n(n+1)/2$). This degeneracy is a property of the Coulomb force, and is related to the separation of the differential equation in parabolic as well as spherical coordinates. Parabolic coordinates are especially useful in treating the positive energy problem of Coulomb scattering. There is also another observable that commutes with the Hamiltonian H called the Lenz vector:

$$\vec{M} = \frac{1}{2m}(\vec{p}\times\vec{L} - \vec{L}\times\vec{p}) - e^2\frac{\vec{r}}{r}; \quad [\vec{M}, H] = 0. \tag{9.3}$$

The Lenz vector operator contains matrix elements of $1/r$, and can be used to solve the hydrogen atom by algebraic means, as shown in Landau and Lifshitz's book on *Quantum Mechanics* and in Merzbacher Problem 16.11. A Hamiltonian with a spherically symmetrical potential $V(r)$ different from $V \sim 1/r$ would not commute with $\vec{M}$, and the energy levels would be l dependent. The classical Lenz vector is fixed in space in the plane of the orbit and directed parallel to the semi-major axis of the ellipse. See Problem 9.1. The angular equation is

$$\begin{aligned} &\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + l(l+1)\right) Y_l^{m_l}(\theta,\phi) = 0; \\ &l = 0,1,2,\ldots;\ -l \le m_l \le +l. \end{aligned} \tag{9.4}$$

The first three solutions to the radial equation are

$$\begin{aligned} R_{10}(r) &= (Z/a_0)^{3/2} 2e^{-Zr/a_0}, \\ R_{20}(r) &= (Z/2a_0)^{3/2}\left(2 - \frac{Zr}{a_0}\right) e^{-Zr/(2a_0)}, \end{aligned}$$

and

$$R_{21}(r) = (Z/2a_0)^{3/2} \frac{Zr}{a_0\sqrt{3}} e^{-Zr/2a_0}. \tag{9.5}$$

Note that the exponential in the radial wave function depends on the principle quantum number n, and that for $l \neq 0, R_{nl}(0) = 0$. For arbitrary $l \leq n-1$:

$$\lim_{r \to 0} R_{nl}(r) \sim r^l. \tag{9.6}$$

This is a consequence of the centrifugal repulsion of finite angular momentum that pushes the wave function away from the origin. The first few spherical harmonics are

$$Y_0^0(\theta,\phi) = \frac{1}{\sqrt{4\pi}}, \quad Y_1^0(\theta,\phi) = \sqrt{\frac{3}{4\pi}} \cos\theta,$$
$$Y_1^{\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin\theta e^{\pm i\phi}. \tag{9.7}$$

The general solution to the hydrogen atom is a product of the radial wave function times a spherical harmonic, and the radial wave function is a power of r times an exponential times an associated Laguerre polynomial:

$$\psi_{nlm}(r,\theta,\phi) = R_{nl}(r) Y_l^{m_l}(\theta,\phi); \quad R_{nl}(r) = C_{nl} e^{-\rho/2} \rho^l L_{n+l}^{2l+1}(\rho);$$
$$\rho = \frac{2Z}{na_0} r; \; a_0 = \frac{\hbar^2}{me^2}. \tag{9.8}$$

The associated Laguerre polynomial is defined as a derivative of the Laguerre polynomial:

$$L_q^p(\rho) = \frac{d^p}{d\rho^p} L_q(\rho); \quad \rho L_q'' + (1-\rho) L_q' + q L_q = 0. \tag{9.9}$$

The three-dimensional symmetric harmonic oscillator, with $V(r) = kr^2/2$ can be separated in both Cartesian and spherical coordinates, In Cartesian coordinates, the problem separates into three one-dimensional harmonic oscillators. The wave function is a product of three familiar wave functions:

$$H = H_x + H_y + H_z, \quad H_x = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2,$$
$$\psi_n(x,y,z) = \psi_{n_x}(x)\psi_{n_y}(y)\psi_{n_z}(z),$$
$$E_n = \hbar\omega_0(n_x + n_y + n_z + \frac{3}{2}), \quad \omega_0 = \sqrt{k/m},$$
$$\psi_{n_x}(x) = C_n H_n(\beta x)e^{-\beta^2 x^2/2}, \quad \beta^4 = \frac{mk}{\hbar^2}. \tag{9.10}$$

$H_n(\xi)$ is the Hermite polynomial satisfying $H_n''(\xi) - 2\xi H_n'(\xi) + 2nH_n(\xi) = 0$. Hermite polynomials have the parity of n: $H_n(-\xi) = (-1)^n H_n(\xi)$. The Schrödinger equation can also be separated in spherical coordinates. The mapping of the various states from one coordinate system to the other is subtle, but can be understood from degeneracy and parity. The ground state is non-degenerate, even parity, $n_x = n_y = n_z = 0$. The spherical solutions have the parity of l, so this state is assigned $n = 1, l = 0$. The first excited state is triply degenerate with odd parity, and is assigned $n = 1, l = 1$. The second excited state $n_x + n_y + n_z = 2$ has six-fold degeneracy and even parity, corresponding to $l = 2, n = 1$ and $l = 0, n = 2$. The energy levels in spherical coordinates are $E_{nl} = \hbar\omega_0(2n + l - 1/2)$. It is interesting to note that solutions to the radial wave equation for the harmonic oscillator in spherical coordinates contain associated Laguerre polynomials, just like the hydrogen atom. The exponential becomes a Gaussian, but otherwise the two solutions are related. See Problem 9.2. So we thought that we had at least two problems in wave mechanics that we could solve exactly, the harmonic oscillator and the hydrogen atom. But in reality, with judicious choice of variables, they are the same problem! Remember how the angular momentum can be generated from two independent one-dimensional harmonic oscillators. Small world.

One way of displaying the energy levels of hydrogen is shown in Figure 9.1, where the first column is all $l = 0$, the second is $l = 1$, etc. In the absence of relativistic corrections the energy levels are the same as l increases for a given n. In what is sometimes called spectroscopic notation, the levels are labeled $n^{2S+1}L_J$, but for hydrogen,

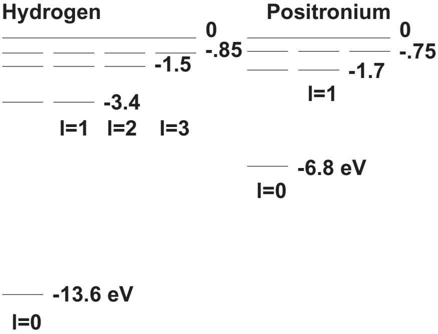


Figure 9.1: Energy levels of hydrogen on the left and positronium on the right, according to the Bohr model. The energy levels are degenerate in the orbital quantum number $0 \le l \le n-1$. Because the postronium reduced mass $m = m_e/2$, the levels are half those of hydrogen for the same principal quantum number n.

neglecting the hyperfine structure which comes from the magnetic moment of the proton, the spin is always 1/2, so we abbreviate the label to nL_J. Then capital letters are used to specify the orbital angular momentum, with the choice that everyone uses beginning with the letter S $(L = 0)$, P $(L = 1)$, D $(L = 2)$, and F $(L = 3)$. For $L = 4$ and beyond it follows the alphabet, but for our purposes we almost never go past $L = 2$. Thus one of our levels would be specified as $2P_{3/2}$. The fine structure was obtained in Chapter 2, Equation (2.155):

$$E_{n,j} = mc^2 + E_0 + \frac{2E_0^2}{mc^2}\left(\frac{3}{4} - \frac{n}{j+1/2}\right), \quad j = l \pm 1/2. \tag{9.11}$$

The constant multiplying the fine structure is $2E_0^2/(mc^2)$ for $n = 2$ becomes $\alpha^2 \times (13.6/16)\,\text{eV} = 4.5 \times 10^{-5}\,\text{eV}$. This is the splitting $2P_{3/2} - 2P_{1/2}$ of frequency $\nu = \Delta E/(2\pi\hbar) = 1.09 \times 10^4\,\text{MHz}$.

The hyperfine splitting for hydrogen is smaller than the fine structure by the ratio m_e/m_p, because the magnetic moment of the proton is inversely proportional to its mass, while in electron spin–orbit coupling the two magnetic moments have the same magnitude. The proton and electron both have spin 1/2, so there is a singlet–triplet splitting that is proportional to the product of the two magnetic moments. The calculation, discussed for positronium in Problem 9.4, is a bit delicate, partly because of a singularity. The formula for S waves only was first derived by Fermi:

$$\Delta E = \frac{8}{3}\alpha^4 2.79 \frac{m_e}{m_p}\frac{m_e c^2}{n^3\hbar^2}\vec{S}_1 \cdot \vec{S}_2,$$

$$\vec{S}_1 \cdot \vec{S}_2/\hbar^2 = \frac{(S(S+1) - 3/2)}{2} = -3/4, S = 0; \quad +1/4, S = 1. \tag{9.12}$$

The factor 2.79 is $g/2$ for the proton. The wave length for the triplet–singlet transition is 21 cm, or a radio frequency of 1420 MHz, which is of great importance in radio astronomy. It has been very precisely measured to 12 decimal places.

9.2. Positronium

Positronium is a bound state of e^-e^+. The binding force is the same as hydrogen, so the solution to the Schrödinger equation is identical, except for the reduced mass, which in the case of positronium is $m = m_e/2$, very nearly 1/2 of the mass term in the hydrogen atom. Since the Bohr radius is inversely proportional to the reduced mass, for positronium $a_0 = 0.1058$ nm. Curiously enough, however, the two atoms are the same size, as shown in Figure 9.2. This is because the radial coordinate connects the two bodies, and in positronium the electron and positron chase each other around a circle of diameter

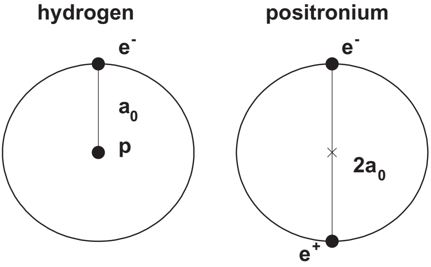


Figure 9.2: Classical circular orbits for hydrogen and positronium. The Bohr radius is inversely proportional to the reduced mass, but the two atoms are the same size.

0.1058 nm, with the center of mass in the middle. For hydrogen, on the other hand, the Bohr radius is 0.0529 nm, but it is measured from the center of mass, where the proton sits. The energy levels in the Bohr formula are proportional to the reduced mass, so −13.6 eV becomes −6.8 eV for positronium. Changing the energy scale by a factor of two gives the positronium levels in Figure 9.1. So far so good. The fine structure of positronium differs from hydrogen in that the spin–spin interaction, Equation (9.12), is the same magnitude as the $\vec{L} \cdot \vec{S}$ term. For positronium the hyperfine splitting of the ground state is

$$\begin{aligned} E(n = 1,^3 S) - E(n = 1,^1 S) &= \frac{7\alpha^4}{12} m_e c^2 = 8.46 \times 10^{-4}\,\text{eV}, \\ \nu &= 2.046 \times 10^5\,\text{MHz}. \end{aligned} \tag{9.13}$$

The numerical factor 7/6 comes from two terms. One (2/3) is the usual dipole–dipole magnetic interaction that occurs in hydrogen. The other factor of 1/2 comes from the annihilation diagram that returns the $e^- e^+$ pair, and contributes to Bhabha scattering: $e^- e^+ \to \gamma \to e^- e^+$. The annihilation diagram only effects the ^3S state. See Problems 9.3 and 9.4. A careful treatment including various radiative corrections lowers this frequency to $\nu = 2.0338 \pm 0.0004$ MHz, a number that has been confirmed by experiment. Equation (9.13) is larger than the fine structure splitting in hydrogen, Equation (9.11), so both terms are combined into positronium 'fine structure', which gives a singlet–triplet aspect to the energy levels. Regarding the spin-orbit force, while there is the Thomas precession that is a Lorentz transformation effect, there is no orbital magnetic moment in positronium, because there is no net current flow. The total spin has no magnetic moment either, so there is no linear Zeeman effect. See Problem 9.5.

The positronium atom is not stable. If energy conservation allowed electron capture in hydrogen: $e^- + p \to n + \nu$, none of us would be here to write these books. Fortunately for us the neutron is too heavy for this capture reaction to go. However, nothing prevents the annihilation $e^- e^+ \to \gamma + \gamma$, giving two back to back 511 keV γ rays, that are used in medical tomography. The lifetime for this annihilation is about 10^{-10} s, and in that time the e^-e^+ pair make about 10^6 orbits, so the atom is well defined. But annihilation ends it all. Annihilation only occurs in S states, because only those states are finite at $r = 0$, where the pair meets, and annihilation strongly favors $n = 1, l = 0$. There are two spin states: 1^3S_1, and 1^1S_0. Only the singlet state can decay into two photons; the two photon final state is forbidden for triplet positronium. A particle or system with spin one cannot decay exclusively into two photons, due to an argument first proposed by Landau [1]. The proof is based on failed attempts to construct a vector (unit angular momentum) out of the final state vectors $\vec{k}$, $-\vec{k}, \vec{\epsilon}_1$, and $\vec{\epsilon}_2$ that has Bose symmetry and satisfies $\vec{\epsilon}_1 \cdot \vec{k} = \vec{\epsilon}_2 \cdot \vec{k} = 0$. This list of rejected vectors comes from Nachtmann's book [p. 164].

1 $(\vec{\epsilon}_1 \cdot \vec{\epsilon}_2)\vec{k}$
2 $\vec{\epsilon}_1 \times \vec{\epsilon}_2$
3 $(\vec{\epsilon}_1 \times \vec{\epsilon}_2) \times \vec{k}$
4 $(\vec{\epsilon}_1 \times \vec{\epsilon}_2) \cdot \vec{k}\,\vec{k}$
5 $\vec{\epsilon}_1 \cdot \vec{k}\,\vec{\epsilon}_2 + \vec{\epsilon}_2 \cdot \vec{k}\,\vec{\epsilon}_1$
6 $\vec{\epsilon}_1 \times \vec{k}(\vec{\epsilon}_2 \cdot \vec{k}) + \vec{\epsilon}_2 \times \vec{k}(\vec{\epsilon}_1 \cdot \vec{k})$
7 $(\vec{\epsilon}_1 \cdot \vec{k})(\vec{\epsilon}_2 \cdot \vec{k})\vec{k}.$

Three, five, six and seven vanish by the transversality condition, and the rest do not satisfy Bose symmetry that requires invariance under exchange of $\vec{\epsilon}_1$ and $\vec{\epsilon}_2$ with simultaneous $\vec{k} \to -\vec{k}$. Null arguments always leave a bit of doubt. Maybe we forgot something? But nature obeys the argument. Thus, $\pi^0 \to \gamma + \gamma$ is OK, because the π^0 has spin zero, but ρ^0 with spin one does not decay into two γ's. The spin zero Higgs can decay into two γ's, albeit with a small branching ratio. Perhaps it helps to mention that there is no problem constructing a suitable pseudoscalar $\vec{\epsilon}_1 \times \vec{\epsilon}_2 \cdot \vec{k}$, that appears in the amplitude for singlet positronium annihilation (Equation (9.21)). Note that a pseudoscalar gives photons with crossed polarizations, while a scalar $(\vec{\epsilon}_1 \cdot \vec{\epsilon}_2)(-\vec{k} \cdot \vec{k})$ gives parallel polarizations. With the two γ channel closed to triplet positronium, it has a much longer lifetime, and decays into three γ's. This gives us a handy mnemonic, namely, 'triplet positronium decays into three γ rays'. Unfortunately, singlet positronium does not decay into one γ ray. That would violate the conservation of both space momentum and angular momentum, so there has to be two of them. The correlation of the outgoing photon polarization is a good example of quantum entanglement.

The annihilation of positronium is quantum electrodynamics, while the bound state is just the Schrödinger theory. It is possible to calculate the decay rate with a sort of hybrid calculation that we will now demonstrate. Feynman diagrams involve plane waves, and a bound state is described by space wave function, that can be composed of plane waves by Fourier expansion. In Chapter 2, during discussion of the Dirac equation it was mentioned that relativistic space wave functions that could be interpreted as the location of a particle really did not make sense. However, the electron and positron in the ground state of positronium are non relativistic, so the Fourier expansion can be done, with the Feynman rules applied to non relativistic spinors. Accordingly, we write down the positronium ground state wave function in terms of a Fourier expansion. The non-relativistic space wave function $\psi(\vec{x})$ and its Fourier transform $f(\vec{p})$ are defined as

$$\psi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} f(\vec{p}),$$

$$\int d^3x |\psi(\vec{x})|^2 = \int \frac{d^3p}{(2\pi)^3} |f(\vec{p})|^2 = 1. \tag{9.14}$$

The two annihilation diagrams, that resemble Compton scattering, are shown in Figure 9.3. Although the reaction was not discussed in earlier chapters, in an e^-e^+ storage ring the singlet state can form these two diagrams, giving the cross-section for $e^-e^+ \to \gamma\gamma$ at high energy. Using the Feynman rules, the amplitude is

$$\begin{aligned}\langle\gamma\gamma| T \left|e^-e^+\right\rangle &= e^2 \epsilon^*_\mu(k)\epsilon^*_\nu(k')\bar{v}_{s'}(p') \\ &\quad\times \left\{ \frac{\gamma^\mu(\not{p} - \not{k} + m)\gamma^\nu}{2p\cdot k} + \frac{\gamma^\nu(\not{p} - \not{k}' + m)\gamma^\mu}{2p\cdot k'} \right\} u_s(p).\end{aligned} \tag{9.15}$$

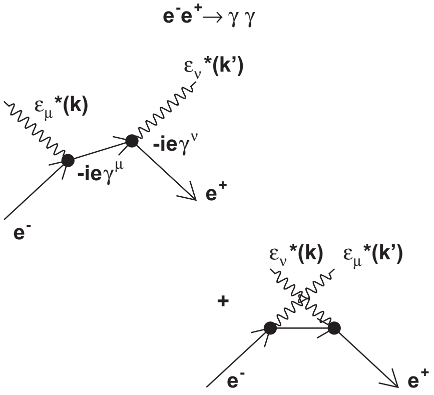


Figure 9.3: Diagrams for the annihilation of singlet positronium into two γ rays. The diagrams are similar to Compton scattering, and the two amplitudes are added together.

You may recall that Compton scattering is a fairly difficult calculation, so before we start we should simplify this expression as much as possible using the fact that the e^-e^+ are very slow. The spinors are given in Chapter 2, Equations (2.121) and (2.122) for $u_s(p)$ and Equation (2.160) for $v_s(p)$. The top two components are large in the non-relativistic limit for $u_s(p)$, and the bottom two components are large for $v_s(p)$, so the operators inside the curly brackets have to lower the two upper components of u down two steps in order to multiply by $\bar{v}$. Also $\not p = m\gamma^0$, and the photon polarization vectors have only space components, so that we may write:

$$\begin{aligned}&\bar{v}_{s'}(p')(\not\epsilon_1^*(k)(\not p - \not k + m)\not\epsilon_2^*(k'))u_s(p)\\&\quad = \bar{v}_{s'}(p')(m\not\epsilon_1^*\not\epsilon_2^*(-\gamma^0 + 1) - \not\epsilon_1^*\not k\not\epsilon_2^*)u_s(p). \qquad (9.16)\end{aligned}$$

The matrix $-\gamma^0+1$ vanishes operating on the upper two components of $u_s(p)$, and the denominators are the same $2p \cdot k = 2p \cdot k' = 2m^2$, so we may approximate Equation (9.15) as

$$\langle\gamma\gamma|\,T\,\left|e^-e^+\right\rangle = -\frac{e^2}{2m^2}\bar{v}_{s'}(p')(\not\epsilon_1^*\not k\not\epsilon_2^* + \not\epsilon_2^*\not k'\not\epsilon_1^*)u_s(p). \qquad (9.17)$$

Now choose the z axis for the final state back-to-back γ's. Then $k = (m, 0, 0, m)$ and $k' = (m, 0, 0, -m)$ with $\epsilon_1 = (0, \epsilon_{1x}, \epsilon_{1y}, 0)$, and

$\epsilon_2 = (0, \epsilon_{2x}, \epsilon_{2y}, 0)$, so that the four vectors and the space vectors are orthogonal: $\epsilon_1 \cdot k = \epsilon_2 \cdot k' = 0$, and $\vec{\epsilon}_1 \cdot \vec{k} = \vec{\epsilon}_2 \cdot \vec{k}' = 0$. Using these choices, one may reduce Equation (9.17):

$$\begin{aligned}&\bar{v}_{s'}(p')(\not\epsilon_1^* \not k \not\epsilon_2^* + \not\epsilon_2^* \not k' \not\epsilon_1^*) u_s(p)\\ &\quad = \bar{v}_{s'}(p')((\not k' - \not k)\not\epsilon_1^* \not\epsilon_2^* - 2\epsilon_1^* \cdot \epsilon_2^* \not k') u_s(p),\end{aligned} \tag{9.18}$$

and

$$\begin{aligned}\not\epsilon_1^* \not\epsilon_2^* &= -\vec{\epsilon}_1^{\,*} \cdot \vec{\epsilon}_2^{\,*} + \gamma^1\gamma^2(\epsilon_{1x}^*\epsilon_{2y}^* - \epsilon_{2x}^*\epsilon_{1y}^*)\\ &= -\vec{\epsilon}_1^{\,*} \cdot \vec{\epsilon}_2^{\,*} + \gamma^1\gamma^2 \frac{\vec{\epsilon}_1^{\,*} \times \vec{\epsilon}_2^{\,*} \cdot \vec{k}}{m}.\end{aligned} \tag{9.19}$$

Using $\not k = m(\gamma^0 - \gamma^3), \not k' = m(\gamma^0 + \gamma^3)$, and the explicit forms for the γ matrices and the spinors (normalization gives $2m$) results in the following form for Equation (9.17):

$$\begin{aligned}\langle\gamma\gamma| T \left|e^- e^+\right\rangle &= -\frac{e^2}{2m^2} 4m^2 \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}\\ &\times \begin{pmatrix} \vec{\epsilon}_1^{\,*} \cdot \vec{\epsilon}_2^{\,*} & 0 & i\vec{\epsilon}_1^{\,*} \times \vec{\epsilon}_2^{\,*} \cdot \vec{k}/m & 0 \\ 0 & \vec{\epsilon}_1^{\,*} \cdot \vec{\epsilon}_2^{\,*} & 0 & i\vec{\epsilon}_1^{\,*} \times \vec{\epsilon}_2^{\,*} \cdot \vec{k}/m \\ -i\vec{\epsilon}_1^{\,*} \times \vec{\epsilon}_2^{\,*} \cdot \vec{k}/m & 0 & -\vec{\epsilon}_1^{\,*} \cdot \vec{\epsilon}_2^{\,*} & 0 \\ 0 & -i\vec{\epsilon}_1^{\,*} \times \vec{\epsilon}_2^{\,*} \cdot \vec{k}/m & 0 & -\vec{\epsilon}_1^{\,*} \cdot \vec{\epsilon}_2^{\,*} \end{pmatrix}\\ &\times \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.\end{aligned} \tag{9.20}$$

Note that there is no spin flip operator in the matrix. We have chosen opposite spins $|{\uparrow\downarrow}\rangle$ that give a non-zero result. Remember that $\bar{v}_{s'}(p')$ is a positron spinor, that has the spin states flipped. Multiplying it all out gives:

$$\langle\gamma\gamma| T \left|e^- e^+\right\rangle = -2ie^2 \frac{\vec{\epsilon}_1 \times \vec{\epsilon}_2 \cdot \vec{k}}{m} |{\uparrow\downarrow}\rangle. \tag{9.21}$$

QED automatically satisfies Landau's rule that $m_s = 0$. The two photon polarization vectors are orthogonal, consistent with total angular momentum zero $(\epsilon_{1x}\epsilon_{2y} - \epsilon_{2x}\epsilon_{1y})/2 = i(\vec{\epsilon}_1^{\,*\lambda=+1} \times \vec{\epsilon}_2^{\,*\lambda=-1})$.

This correlation is the quantum entanglement mentioned earlier. The two choices for initial spins combine to form a singlet state:

$$\langle\gamma\gamma|\, T \left|e^-e^+; S=0\right\rangle = -\sqrt{2}ie^2\vec{\epsilon}_1 \times \vec{\epsilon}_2 \cdot \hat{k}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \tag{9.22}$$

This expression gets folded in to a Fourier expansion. It is independent of the relative momentum of the e^-e^+ pair, a negligibly small number compared to the γ energies, so it serves as a multiplier of the integral. The next task is to get the normalization right. Our spinors are normalized to $2E$, which is $2m$ in this case, where m is the electron mass, and ordinary kets in quantum mechanics are normalized to one, so following Peskin and Schroeder, p. 149, we write the bound state ket as

$$|B\rangle = \sqrt{4m}\int \frac{d^3p}{(2\pi)^3} f(p)\frac{1}{\sqrt{2m}}\frac{1}{\sqrt{2m}}\frac{8\pi i\alpha\vec{\epsilon}_1 \times \vec{\epsilon}_2 \cdot \hat{k}}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \tag{9.23}$$

The multiplier $\sqrt{4m}$ in front changes the state back to our normalization, so it can be used in our usual formulas. The integral is just $\psi(0)$:

$$|B\rangle = \psi(0)\frac{1}{\sqrt{m}}\frac{8\pi i\alpha\vec{\epsilon}_1 \times \vec{\epsilon}_2 \cdot \hat{k}}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \tag{9.24}$$

The sum over photon polarization states gives a factor of two:

$$\sum_{\text{spins}} \langle B|B\rangle = \frac{128\pi^2\alpha^2}{m}|\psi(0)|^2. \tag{9.25}$$

This expression is ready to go into our regular formula for the decay rate of an unstable particle, in this case positronium with mass $2m$. So far we are working in natural units, and we will convert $\psi(0)$ into these units presently. Note that this derivation makes no assumptions about exactly what $\psi(0)$ is. The factor of α^2 comes from the annihilation. So, this analysis could apply to any bound S state that decays into two photons.

$$d\Gamma = \frac{1}{4m}\frac{(2\pi)^4}{(2\pi)^6}\frac{128\pi^2\alpha^2}{m}|\psi(0)|^2\delta(\omega+\omega'-2m)\delta(\vec{k}+\vec{k}\,')\frac{d^3k}{2m}\frac{d^3k'}{2m}. \tag{9.26}$$

Performing the integration and multiplying by 4π for the solid angle gives:

$$\Gamma(e^-e^+\ 1^1S_0 \to \gamma\gamma) = \frac{4\pi\alpha^2}{m^2}|\psi(0)|^2. \tag{9.27}$$

For hydrogen, $\psi(0) = (1/\sqrt{\pi})(1/a_0)^{3/2}$, where $a_0 = \hbar^2/(me^2)$ is the Bohr radius, which gets doubled for positronium. So, in terms of α in natural units, the positronium Bohr radius is $2a_0 = 2/(\alpha m)$. Substituting all of this into Equation (9.27) gives the decay rate for positronium in eV, converted to s^{-1} by dividing by $\hbar = 6.58 \times 10^{-16}$ eV-s:

$$\begin{aligned}\Gamma(e^-e^+\ 1^1S_0 \to \gamma\gamma) &= \frac{\alpha^5\, m}{2} = 5.3 \times 10^{-6}\,\text{eV}\\ &= 0.8 \times 10^{10}\,\text{s}^{-1},\\ \tau = 1/\Gamma &= 1.25 \times 10^{-10}\,\text{s}.\end{aligned} \tag{9.28}$$

The experimental number for the decay rate is $\Gamma = 7.99 \pm 0.11 \times 10^9\,\text{s}^{-1}$ [2] in good agreement with the calculation.

The triplet state decays into three γ's, a more complicated calculation in QED. The theoretical number, which agrees with experiment, is

$$\Gamma(1^3S_1 \to 3\gamma) = \frac{2(\pi^2 - 9)}{9\pi}\alpha^6 m = 7.2 \times 10^6\,\text{s}^{-1}. \tag{9.29}$$

The triplet state lives about 1000 times longer than the singlet state.

9.3. Other Exotic Coulomb Atoms

Positronium is not the only exotic atom bound by the Coulomb force that has been seen experimentally. Muonium, a bound state of μ^+e^-, has been formed by stopping positive muons in a suitable gas, and looking for characteristic precession of the atom in a magnetic field. The atom disappears when the muon decays $\mu^+ \to e^+\nu_e\bar{\nu}_\mu$ with a lifetime of $2.2\,\mu$s. The decay positron follows the muon spin, that can precess in a magnetic field. The muon Larmor frequency is about $\omega_\mu = 8.4 \times 10^8$ rad/s/T, so in a modest magnetic field like 100 gauss (10^{-2} T) the muon precesses through 2π in $0.75\,\mu$s, or about three turns in one lifetime. However, triplet mounium couples the muon spin to the electron spin, and the electron magnetic moment is 200 times larger, so one of the signatures of muonium formation is much faster precession of the decay positrons. See Problem 9.6. Muonium is actually very similar to hydrogen, since the muon mass is 200 times the electron mass. The discovery of muonium formation in argon gas was done by V. W. Hughes *et al.* [3]. The lifetime is long enough for muonium to form molecules with ordinary atoms, and some chemical binding studies have been made. The μ^+e^- system could also annihilate into two neutrinos, a reaction very similar to positronium annihilation, only going through the weak force. A dimensional argument would lower the rate relative to normal decay by about 10^{-12} because $|\psi(0)|^2$ would replace m_μ^5 by $m_\mu^2 m_e^3 \alpha^3$. See Problem 9.7. The author is not aware of this process ever being observed.

There also are $\pi - \mu$ atoms that have been formed in the three body decay $K_L^0 \rightarrow \pi^\pm \mu^\mp \nu$ [4]. The pion and muon are collinear on the phase space boundary, and there is a region where the relative velocity is very low, so that the atom can form. The neutral atom travels in a straight line to a thin foil, where it is ionized and the π and μ are separately detected. The measured formation rate was consistent with expectations for a coulomb attractive force with no anomalies. While muonium continues to be studied, this is the only instance of $\pi - \mu$ atoms that the author is aware of. Other exotic atoms of interest are formed by $q\bar{q}$ pairs. The quarks are charged, but they are bound by the strong force, which gives extra meaning to the term ‘exotic’.

9.4. Bound States of $Q\bar{Q}$ Pairs

The existence of heavy quarks makes the quark model real. The assignment of a mass to the light quarks, one of the defining properties of a particle, is complicated by their confinement. For the heavy quarks (c, b) the binding energies into D (1.9 GeV mass) and B (5 GeV mass) mesons are smaller fractions of the quark mass, so, within certain limitations, the masses of these heavy quarks are easier to define. We have seen in Chapter 8 how this simplification breaks down for the 125 GeV Higgs, but let us treat that as a special case. Another special case is the 175 GeV top quark, that lives too short a time to form mesons, and decays directly into Wb, so that its mass can be reconstructed from its decay products. Among the six known quarks, only the top decays as a free particle. The other five decay within bound states of $q\bar{q}$ mesons, or qqq baryons. Most D and B mesons contain only one heavy quark, although the $B_c^+ = c\bar{b}$ with two heavy quarks is now being studied at LHCb. Baryons with two heavy quarks should exist, and candidates with two charmed quarks like $\Xi_{cc}^+ = (c, c, d)$ and $\Xi_{cc}^{++} = (c, c, u)$ have been reported. Mesons and baryons with more than one heavy quark are very useful for understanding quark binding, because they contain a heavy di-quark — something new.

Discovery channel examples of pairs of heavy quarks are $(c\bar{c})$ charmonium and $(b\bar{b})$ bottomonium, the analogs of positronium for a strongly bound $q\bar{q}$ pair. The charmonium ground state, the J/ψ at 3.1 GeV was introduced in Chapter 1. All quarkonium states decay by strong interaction, but the low lying states have narrow widths because the decays are strongly inhibited. For example, the J/ψ width $\Gamma = 87\,\text{keV}$, for a lifetime of about 10^{-20} s, 10 orders of magnitude shorter than positronium. However, the atomic radius for charmonium is about 10^{-4} times that of positronium, and the velocity in the orbit is about 50 times higher, giving perhaps 1000 orbits of the atom before it disappears, long enough for an application of the Schrödinger equation to the bound state energy levels.

The J/ψ was discovered independently by the SPEAR e^-e^+ storage ring at SLAC, and at a fixed target experiment at Brookhaven where the invariant masses of e^-e^+ pairs were measured in $p + \text{Be} \to e^-e^+ + \text{anything}$. The three lowest lying $b\bar{b}$ bottomonium states were discovered in a fixed target experiment at Fermilab, where the masses of $\mu^-\mu^+$ pairs were measured after 400 GeV protons struck a copper or platinum target. In their first report, two of the resonances at 9.44 ± 0.03 and 10.17 ± 0.05 GeV were resolved [5]. The third one was found later by the same group. At the same time at Cornell a new e^-e^+ storage ring named CESR and pronounced ‘Caesar’ was under construction, with a maximum ring energy of 6 GeV. Operations

Table 9.1: Bottomonium energy levels.

State	J^{PC}	Mass (MeV)	Γ	Decay mode
$\Upsilon(4S)$	1^{--}	10579	20 MeV	$B\bar{B}$
$\chi_{b2}(3P)$	2^{++}	10524		
$\chi_{b1}(3P)$	1^{++}	10513		
$\Upsilon(3S)$	1^{--}	10355	20 keV	$\Upsilon(2S)+X$
$\chi_{b2}(2P)$	2^{++}	10268		$\Upsilon(2S)\gamma+\Upsilon(1S)\gamma$
$\chi_{b1}(2P)$	1^{++}	10255		$\Upsilon(2S)\gamma+\Upsilon(1S)\gamma$
$\chi_{b0}(2P)$	0^{++}	10232		$\Upsilon(2S)\gamma+\Upsilon(1S)\gamma$
$\Upsilon(2S)$	1^{--}	10023	32 keV	$\Upsilon(1S)\pi\pi$, $\chi\gamma$, $e^-e^+2\%$
$\chi_{b2}(1P)$	2^{++}	9912		$\Upsilon(1S)\gamma$
$\chi_{b1}(1P)$	1^{++}	9893		$\Upsilon(1S)\gamma$
$\chi_{b0}(1P)$	0^{++}	9859		$\Upsilon(1S)\gamma$
$\Upsilon(1S)$	1^{--}	9460	54 keV	$ggg \to$ hadrons, $e^-e^+2.4\%$
$\eta_b(1S)$	0^{-+}	9399		analog of ${}^1S_0$ positronium

began in 1979 at about 5 GeV energy — the bottomonium region, and operated at those energies for 20 years. It was a fine example of synergy between different techniques to study the same phenomenon. Fermilab discovered the resonances, but Cornell could study them thoroughly, and also could use the $\Upsilon 4S$ to learn about B mesons that contained the new quark. Figures 9.5 and 9.6 show CESR data used to unravel the atomic energy level scheme shown in Figure 9.4. Some of the energy levels are given in Table 9.1.

$\Upsilon(1S), (2S)$, and $(3S)$ states have very narrow widths. $\Upsilon(1S)$ decays into hadrons via a three gluon intermediate state. Landau's theorem applies equally well to photons and massless spin one gluons, even though all gluons are confined, so $\Upsilon \to gg$ is forbidden. The γ energies in Figure 9.6 give the level spacing between $\Upsilon(2S)$ and the $\chi_b(1P)$ states. Check the arithmetic! The χ states are all triplet P(0,1,2), and are created by Υ decay. Electric dipole is favored in the single γ decays, because the 'atom' is small compared to the radiated wave length. The lowest lying level is the $\eta_b(1S)$ at 9399 MeV, that is a singlet state, with quantum numbers 0^{-+}. This state could decay into two γ's, but that mode has not been seen to date (PDG). The branching ratios for $\Upsilon(1S) \to e^-e^+ = 2.4\%$ and $\Upsilon(2S) = 2\%$

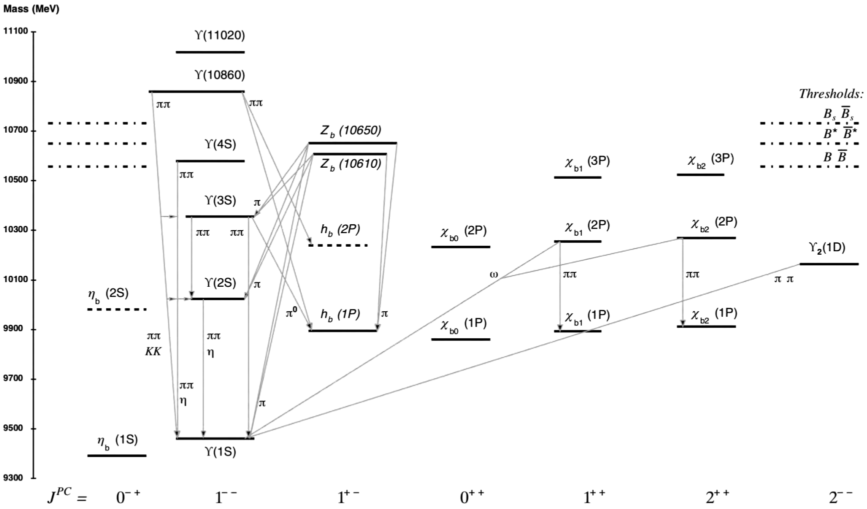

Mass (MeV)
11100
10900
10700
10500
10300
10100
9900
9700
9500
9300
Υ(11020)
Υ(10860)
Υ(4S)
Υ(3S)
Υ(2S)
Υ(1S)
η_b (2S)
η_b (1S)
ππ
ππ
ππ
ππ
ππ
ππ
ππ
KK
ππ
η
ππ
η
π
π
π
π^0
π
ω
ππ
ππ
π π
Z_b (10650)
Z_b (10610)
h_b (2P)
h_b (1P)
χ_{b0} (2P)
χ_{b0} (1P)
χ_{b1} (3P)
χ_{b1} (2P)
χ_{b1} (1P)
χ_{b2} (3P)
χ_{b2} (2P)
χ_{b2} (1P)
Υ_2(1D)
Thresholds:
$B_s \bar{B}_s$
$B^* \bar{B}^*$
$B \bar{B}$
J^{PC} =
0^{-+}
1^{--}
1^{+-}
0^{++}
1^{++}
2^{++}
2^{--}

Figure 9.4: Energy levels of the bottomonium atom. The quantum numbers of the levels are shown along the x-axis as J^{PC}. The 1^{--} column corresponds to the peaks in Figure 9.5. The gamma rays in Figure 9.6 are from the $\Upsilon(2S)$ to the three $1P$ states $0^{++}, 1^{++}$, and 2^{++}. The Z states with 1^{+-} decay into $\Upsilon\ \pi$, so contain $b\bar{b}$, but do not have quantum numbers of a pure $q\bar{q}$ state, and are 'exotic'. Only the 1^{--} states can be created directly by e^-e^+ annihilation. Credit Particle Data Group.

are shown, because these rates can be calculated in QED in analogy with $e^-e^+ \to \gamma\gamma$:

$$\Gamma(^3S_1 \to l^-l^+) = \frac{16\pi\alpha^2 e_q^2}{M^2}|\psi(0)|^2, \tag{9.30}$$

where e_q is the fractional quark charge ($-1/3$ for b quarks), and M is the mass of the vector meson. Equation (9.27) for positronium decay has the electron mass squared in the denominator. Multiplying and dividing by four to convert the mass to positronium results in the same equation as (9.30), except for the quark charge. We shall see that some knowledge of the wave function can be obtained from fits to the energy levels, so that $|e_q| = 1/3$ can be confirmed from bottomonium data. For the ^3S levels we have:

$$\Upsilon(1S) \to \Upsilon(2S) = 563\,\text{MeV}, \quad \text{and} \quad \Upsilon(1S) \to \Upsilon(3S) = 895\,\text{MeV}. \tag{9.31}$$

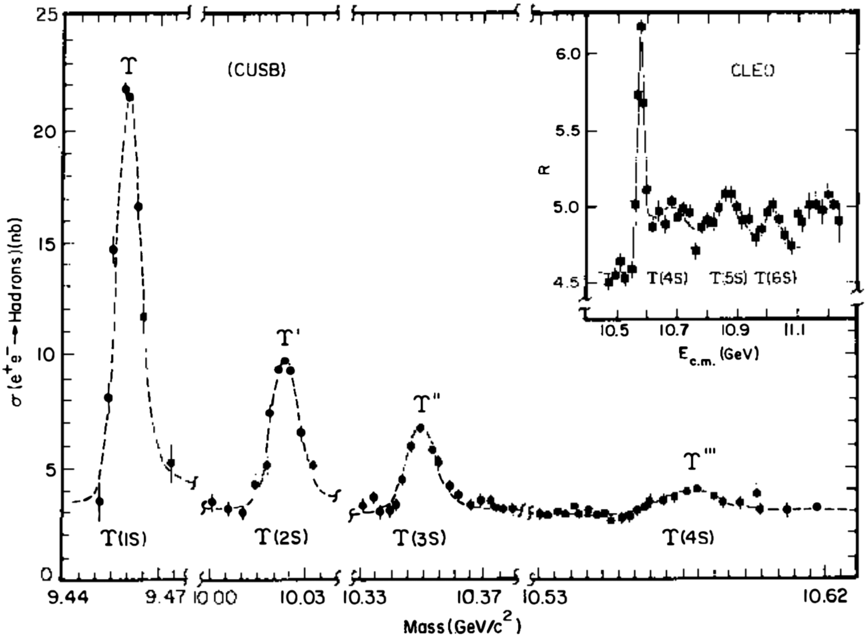


Figure 9.5: Cross-section for $e^-e^+ \rightarrow$ hadrons at the Cornell storage ring CESR. Four peaks of $\Upsilon\ ^3S$ resonances are clearly seen. Note that the energy scale is fragmented to show the peak regions. The $\Upsilon(4S)$ is broader because it has enough energy to decay into $B\bar{B}$ mesons, and is a copious source of these interesting particles. Credit Ref. [14].

Figure 9.5 shows the $\Upsilon(1S)$, $(2S)$, $(3S)$, $(4S)$ resonances in an energy scan of the CESR e^-e^+ storage ring at Cornell. The $\Upsilon(4S)$ is above threshold for $B\bar{B}$ meson decay. Higher resonances are shown in the inset.

Figure 9.6 shows the single γ energy spectrum recorded from running the CESR storage ring on the $\Upsilon(2S)$ at 10,023 MeV. The energies of the $1P$ levels that lie below the $\Upsilon(2S)$ are determined from the energies of these three peaks. These electric dipole transitions feed the $n = 1$, $l = 1$ levels that are not accessible directly by e^-e^+ anihilation.

The companion table for charmonium is Table 9.2.

When compared to bottomonium, the charmonium spectrum is truncated, missing three χ states and one long lived $\psi(3S)$, because

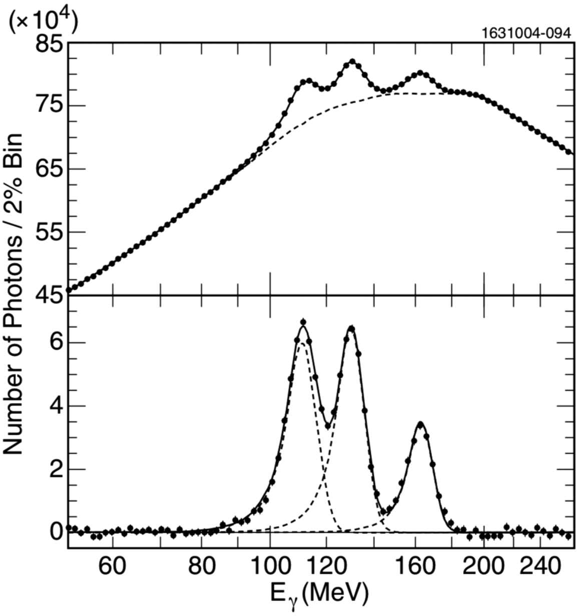


Figure 9.6: CLEO Collaboration single γ energy spectrum running on the $\Upsilon(2S)$. The smooth background is from $\pi^0 \to \gamma\gamma$ decay. The three γ peaks at approximately 110, 130, and 160 MeV feed the $1P$ states with $J = 2, 1, 0$. Credit Ref. [15].

Table 9.2: Charmonium energy levels.

State	J^{PC}	Mass (MeV)	Γ	Decay mode
$\psi(4040)$	1^{--}	4039	80 MeV	$D\bar{D}$
$\psi(3770)$	1^{--}	3774	27 MeV	$D\bar{D}$
$\psi(2S)$	1^{--}	3686	294 keV	$J/\psi(1S)\pi\pi\ \chi\gamma$
$\chi_{c2}(1P)$	2^{++}	3556	2.0 MeV	$J/\psi(1S)\gamma$
$\chi_{c1}(1P)$	1^{++}	3511	0.85 MeV	$J/\psi(1S)\gamma$
$\chi_{c0}(1P)$	0^{++}	3415	10 MeV	$J/\psi(1S)\gamma$
$J/\psi(1S)$	1^{--}	3097	93 keV	$ggg \to$ hadrons, e^-e^+ 6%
$\eta_c(1S)$	0^{-+}	2984		analog of $^1S_0$ positronium

the $\psi(3770)$ is above threshold to decay into a $D\bar{D}$ meson pair. Higher states all decay into D mesons. Energy differences between ^3S states are

$$J/\psi \to \psi(2S) = 589\,\text{MeV}, \quad J/\psi \to \psi(4040) = 943\,\text{MeV}. \quad (9.32)$$

The similarity of Equations (9.31) and (9.32), with the overall energy scale a factor of three larger for bottomonium, has been an important driver in attempts to fit both sets of energy levels with a single potential $V(r)$ in the Schrödinger equation. We shall discuss these attempts in some detail.

The application of the Schrödinger equation to the problem of heavy quarkonium energy levels assumes non-relativistic mechanics. For $c\bar{c}$ the Lorentz factor $\gamma \sim 1.2$, and for $b\bar{b}$ $\gamma \sim 1.1$, so the non-relativistic approximation improves with increasing quark mass. The top quark mass was not known during most of the work on heavy quark potentials, and there was hope that a top quark in the 30 GeV mass range would be an even better subject for the Schrödinger equation. Unfortunately, nature did not cooperate, and a 175-GeV top quark does not live long enough to form bound states, so we have only $c\bar{c}$ and $b\bar{b}$ with which to work. A single potential $V(r)$ assumes that the $q\bar{q}$ force is flavor independent, that is, it is the same for b and c quarks. The quarks have different charges, and that affects their QED transition rates, but the binding force is assumed to be charge independent. This assumption is reasonable if color QCD is the strong binding force, because each quark flavor comes in the same three colors. The decay rates for $J/\psi \to e^-e^+ = 5.6\,\text{keV}$, and $\Upsilon(1S) \to e^-e^+ = 1.3\,\text{keV}$ from the numbers in Tables 9.1 and 9.2, with a ratio 4.3, very close to four expected from e_c^2/e_b^2, implying that $|\psi(0)|^2/M^2$ nearly cancels (Equation (9.30)).

The potential $V(r)$ has to be confining — that is it extends to infinity so that there are no unbound scattering states, otherwise the quarks could exist as free particles. An infinite well, a linear potential (constant force), and a harmonic oscillator are all confining potentials. In one dimension we have Table 9.3

For the infinite well, the integers are $n = 1, 2, 3, 4, \ldots$, and the positive and negative parity solutions have different dependence on n. In three dimensions it is only the odd parity solutions that are allowed. For the linear potential the solution was obtained from the WKB approximation, available in any standard quantum mechanics text. In this approximation the energy levels are given by an integral

Table 9.3: Confining potentials in one dimension.

Potential	E_n	f (mass)	f (strength)
$V = 0 \ -a \leq x \leq +a$ $V = \infty \ \lvert x\rvert > a$	$\hbar^2 k_n^2/2m$ $k_n^+ = (2n+1)\pi/(2a);\ k_n^- = n\pi/a$	m^{-1}	NA
$V = \lambda \lvert x\rvert$	$((3\pi\lambda\hbar(n+1/2))/4\sqrt{2m})^{2/3}$	$m^{-1/3}$	$\lambda^{2/3}$
$V = kx^2/2$	$\hbar\sqrt{k/m}(n+1/2)$	$m^{-1/2}$	$k^{1/2}$

of the momentum between two classical turning points:

$$\int_{x_1}^{x_2} dx(2m(E - V(x)))^{1/2} = \pi\hbar(n + 1/2). \tag{9.33}$$

The integration is elementary. It is a property of the harmonic oscillator that the WKB approximation gives the exact answer for the energy levels, so one may solve the problem either way. The WKB approximation also does a good job compared to the exact solutions to the linear potential [6]. If we are looking for a potential for which the mass dependence of the level spacing is very weak, we see that the linear potential is the best choice. The dependence on strength is a general property of solutions to the Schrödinger equation. If you increase the slope of the potential, the energy levels get farther apart, and if you decrease the slope they get closer together. The spacing of the energy levels of hydrogen is a good example of this behavior.

In passing to three dimensions it is often convenient to substitute $u(r) = rR(r)$ in Equation (9.1). This substitution eliminates the first derivative, and gives an equation that, apart from the centrifugal repulsion, resembles the one-dimensional form:

$$u''(r) + \frac{2m}{\hbar^2}\left(E - V(r) - \frac{\hbar^2 l(l+1)}{2mr^2}\right)u(r) = 0; \quad u(0) = 0. \tag{9.34}$$

The boundary condition at the origin eliminates the even parity solutions to the one dimensional problem, and r is defined as $0 \leq r \leq \infty$.

This modifies the form of the WKB approximation, that becomes:

$$\int_0^{r_c} dr \left(2m\left(E - V(r) - \frac{\hbar^2 l(l+1)}{2mr^2}\right)\right)^{1/2} = \pi\hbar(n + l/2 - 1/4). \tag{9.35}$$

This formula accounts for the χ states with $l = 1$ that lie in between the $^3S_1$ states. Thus, $n = l = 1$ lies between $n = 1, l = 0$ and $n = 2, l = 0$. The splitting of the χ states is due to a spin–orbit force. We are particularly interested in the ^3S states, for which $l = 0$, and the classical turning point is given by $V(r_c) = E$. For $l = 0$ and $V(r) = \lambda r$ the result is

$$\int_0^{r_c} dr(E - \lambda r)^{1/2} = \pi\hbar(n - 1/4)/\sqrt{2m}; \quad \text{gives}$$

$$E_n = (3\pi\lambda\hbar(2n - 1/2)/(4\sqrt{2m}))^{2/3}. \tag{9.36}$$

The even parity states are missing in the three dimensional solution, but otherwise the two energies look the same.

The reason that the logarithmic potential has been a candidate for quarkonium can be shown by using the WKB approximation. For $V(r) = C\ln(r/r_0)$, the potential diverges both for $r \to 0$ and $r \to \infty$, but the slope $\sim 1/r$, so going to infinity the curve is almost flat, very different from a constant force. Also the potential crosses zero at $r = r_0$. The WKB approximation for $l = 0$ reads:

$$\int_0^{r_c} dr(E - C\ln(r/r_0))^{1/2} = \pi\hbar(n - 1/4)/\sqrt{2m}.$$

Let $\xi = r/r_0$, then

$$\int_0^{r_c/r_0} d\xi(E - C\ln\xi)^{1/2} = \frac{\pi\hbar}{r_0\sqrt{2m}}(n - 1/4). \tag{9.37}$$

The definite integral can be evaluated. Let $y = E - C\ln\xi$. Then

$$\xi = e^{(E-y)/C};\ d\xi = -\frac{\xi dy}{C};\ \xi = 0,\ y = \infty;\ \xi = r_c/r_0,\ y = 0;$$

$$\int_0^{r_c/r_0} d\xi(E - C\ln\xi)^{1/2} = \frac{e^{E/C}}{C}\int_0^{\infty} dy\, e^{-y/C}\sqrt{y}. \tag{9.38}$$

The integral is a gamma function: $\Gamma(3/2) = \int_0^\infty dt e^{-t}\sqrt{t} = 1/2\Gamma(1/2) = \sqrt{\pi}/2$. Gamma functions satisfy $\Gamma(x+1) = x\Gamma(x)$, and this property precludes $\Gamma(x)$ from obeying a differential equation. Most special functions obey differential equations. Assembling all of the parts and solving for the energy gives:

$$E_n = C\left\{\ln\left(\frac{\hbar c(n-1/4)}{Cr_0}\right) + \frac{1}{2}\ln\left(\frac{2\pi C}{mc^2}\right)\right\}. \tag{9.39}$$

This result gives the reason for interest in the logarithmic potential as a trial for quarkonium. The mass term is a constant, not multiplying the energy level quantum numbers, so that the level spacing is mass independent. This is the only simple potential that has this property. On the other hand, the confinement is pretty feeble. The force goes to zero at large distances, letting the quarks loose.

The gluons of QCD are massless, so gluon exchange should give a short distance potential that is Coulomb-like. A linear potential is added on for confinement, giving:

$$V(r) = -\frac{4A\alpha_s(m_q)}{3r} + \lambda r. \tag{9.40}$$

The 4/3 is a color factor, that comes from $q\bar{q} \to q\bar{q}$ via one gluon exchange, where the $q\bar{q}$ pair is in a colorless state. This factor is derived in the Appendix. The QCD running coupling constant is evaluated at the quark mass. The formula for α_s was given in Equation (8.87). Setting $\Lambda = 300\,\text{MeV}$, and $m_b = 4.2\,\text{GeV}$ and five flavors gives $\alpha_s(m_b) = 0.31$, and $m_c = 1.4\,\text{GeV}$ and four flavors gives $\alpha_s(m_c) = 0.49$. The coupling constant is dimensionless, but $1/r$ has the dimensions of energy in our units. To convert to energy in GeV and distances in fermis ($1\,\text{f} = 10^{-13}\,\text{cm}$), we have added a constant $A = 0.2\,\text{GeV-f}$ to the 'Coulomb' term. Then the strength λ is given in GeV/f, and r is measured in fermis. Equation (9.40) crosses the axis (the radius r_0 for the logarithmic potential) at

$$\begin{aligned} r_0 = \sqrt{4A\alpha_s/(3\lambda)} = 0.3\,\text{f} \quad \text{for} \quad m_b = 4.2\,\text{GeV} \\ \text{and} \quad \lambda = 1\,\text{GeV/f}. \end{aligned} \tag{9.41}$$

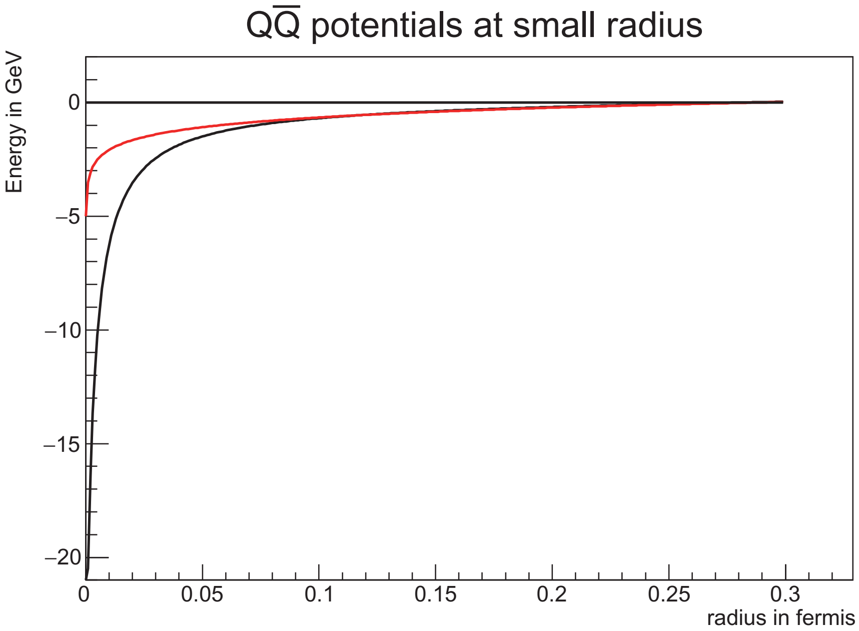


Figure 9.7: $V(r) = -0.083/r + r$ (black) compared to $V(r) = 0.62 \times \log(r/0.29)$ (red) at small r. Both potentials are in GeV units, and vanish at $r = 0.29$ f. Although the two curves agree in the vicinity of the cross-over, the Coulomb-like singularity is much stronger than the logarithm.

For a pure coulomb-like potential, the Bohr radius for a reduced mass $m_b/2$ would be $a_0(b\bar{b}) = \hbar c/(m_b c^2 \times 0.2) = 0.2\,\text{f}$, using $\hbar c = 197\,\text{MeV-f}$, and replacing the fine structure constant $\alpha \to 4\alpha_s/3$. The Bohr radius for the J/ψ is $a_0(c\bar{c}) = 0.4\,\text{f}$. It does not scale like $1/m$ because of the energy dependence of α_s. Figures 9.7 and 9.8 show the QCD inspired potential Equation (9.40) compared to a logarithmic potential that has the same zero, namely $r_0 = 0.29\,\text{f}$, near the origin, and at large distance. The size of the Υ is about $\sqrt{\langle r^2 \rangle} \sim 0.2\,\text{f}$, close to its Bohr radius, and near where these potentials cross the axis. The size of the J/ψ also agrees with the Bohr radius estimate of $\sqrt{\langle r^2 \rangle} \sim 0.4\,\text{f}$. The plots show that the logarithmic potential is a poor match both as $r \to 0$, and $r \to \infty$, but not so bad near the radius of the exotic atom. So, one might expect that the energy levels obtained from the solvable (at least in WKB) logarithmic potential

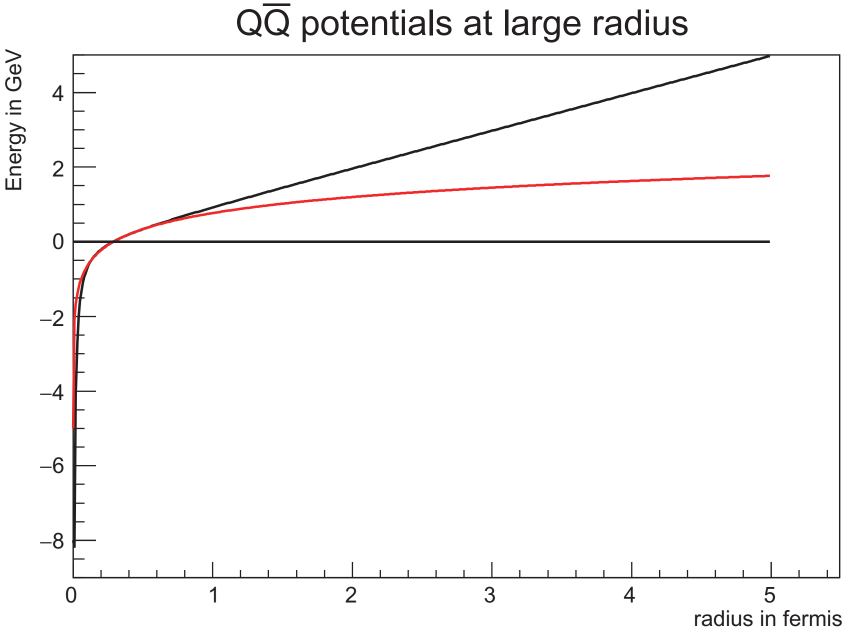


Figure 9.8: Same two potentials as the previous figure for large r. The linear potential is much stronger than the logarithm as $r \to \infty$.

would be satisfactory, but that the wave function at the origin for the logarithmic potential $|\psi(0)|^2$ would not be correct.

The Schrödinger equation with Equation (9.40) for the potential term has to be solved numerically, using techniques that are well developed. We know what each part looks like separately, but we do not have solutions in closed form for the combination. We will not pursue numerical solutions, and satisfy ourselves with some parameters in the linear potential alone. The energy levels for $l = 0$ are given by Equation (9.36). For $l = 1$, the χ states, substitute $2n \to 2n + l$. Working in GeV and fermis.

$$\hbar c = 0.2\,\text{GeV-f}, \quad \lambda_b = 1.8\,\text{GeV/f}, \quad \left\{\frac{3\pi\hbar c\lambda_b}{4\sqrt{m_b c^2}}\right\}^{2/3} = 0.56\,\text{GeV},$$

$$\lambda_c = 1.14\,\text{GeV/f}, \quad \left\{\frac{3\pi\hbar c\lambda_c}{4\sqrt{m_c c^2}}\right\}^{2/3} = 0.59\,\text{GeV}. \tag{9.42}$$

Table 9.4: Linear potential in 3D for quarkonium.

n	l	Level	$b\bar{b}$ model (GeV)	$b\bar{b}$ data (GeV)	$\Delta b\bar{b}$ (MeV)	$c\bar{c}$ model (GeV)	$c\bar{c}$ data (GeV)	$\Delta c\bar{c}$ (MeV)
1	0	(1^3S)	9.460	9.460	0	3.097	3.097	0
1	1	(1^3P)	9.760	9.888	128	3.410	3.494	84
2	0	(2^3S)	10.020	10.023	0	3.686	3.687	0
3	0	(3^3S)	10.470	10.355	−115	4.159	4.039	−120

The masses in the denominators are $m_b = 4.2\,\mathrm{GeV}$ and $m_c = 1.4\,\mathrm{GeV}$. (The formula has $2m$, where m is the reduced mass.) We have adjusted the strength parameter for charmonium to account for the smaller mass. This violates the flavor independence, but this is only an exercise for illustration. See Table 9.4. The logarithmic potential, that is mass independent, is left for the problem set.

The energy for the (1^3P) states is the simple average of the three levels. The strength parameters λ were chosen to match the $(2^3S) - (1^3S)$ energy difference. Differences for the 'predicted' levels are good to about 20%.

The Thomas precession gives an $\vec{L} \cdot \vec{S}$ term of the form:

$$\begin{aligned} -\frac{1}{2m^2c^2r}\frac{dV}{dr}\vec{L}\cdot\vec{S} &= -\frac{\hbar^2}{2m^2c^2r}\frac{dV}{dr}(j(j+1) - l(l+1) - s(s+1))/2 \\ &= +C, j=2; \quad -C, j=1; \quad -2C, j=0; \\ C &= -\frac{\hbar^2}{2m^2c^2}\left\langle \frac{dV}{rdr} \right\rangle. \end{aligned} \tag{9.43}$$

So the ordering of the P states is correctly accounted for by the spin–orbit force if $\langle dV/(rdr)\rangle < 0$, although the magnitudes are not correct, because experimentally the $j = 0$ to $j = 1$ transition is larger than the $j = 1$ to $j = 2$ by about a factor of two. The tensor force operator between two spins is written:

$$S_{12} = \frac{3(\vec{\sigma}_1 \cdot \vec{r})(\vec{\sigma}_2 \cdot \vec{r})}{r^2} - \vec{\sigma}_1 \cdot \vec{\sigma}_2. \tag{9.44}$$

This operator does not depend on j, so one may work in the $|m_l, m_s\rangle$ representation. One of the operators can be expressed in terms of spherical harmonics:

$$\frac{\vec{\sigma} \cdot \vec{r}}{r} = \sqrt{4\pi/3}(\sqrt{2}(Y_1^{-1}\sigma_+ - Y_1^{+1}\sigma_-) + Y_1^0\sigma_z). \tag{9.45}$$

It is then straightforward to work out the matrix elements for ${}^3P$, ${}^3S$, ${}^1P$ and ${}^1S$ states. I got the following results:

$$\sum_{m_l, m_s} |\langle l, s, m_l, m_s| S_{12} |l, s, m_l, m_s\rangle|^2 :$$

$${}^3P = 144/25; \quad {}^1P = 0; \quad {}^3S = 0; \quad {}^1S = 0. \tag{9.46}$$

The matrix element $\langle s = 0, m_s = 0| S_{12} |s = 0, m_s = 0\rangle = 0$, so the tensor force has no effect on singlet spin states. The matrix elements $\langle s = 1, m_s = \pm 1| S_{12} |s = 1, m_s = \pm 1\rangle = 3\cos^2\theta - 1$, can mix $l = 0$ and $l = 2$, but that does not apply to bottomonium.

The tensor force adds to the ${}^3P$, ${}^3S$ splitting, but does not split different j values.

The wave function at the origin, $|\psi(0)|^2$, appears in formulas for ${}^3S_1 \to e^-e^+$, and those decay rates have been measured, so this gives another clue to the correct wave equation. Obviously $|\psi(0)|^2$ is sensitive to the potential near $r = 0$. An equation relating the wave function at the origin to the potential has been attributed to Schwinger [7]:

$$|\psi(0)|^2 = \frac{m}{2\pi\hbar}\left\langle \frac{dV}{dr} \right\rangle, \quad l = 0 \text{ only,}$$

$$|\psi(0)|^2 = \frac{m\lambda}{2\pi\hbar}, \quad \text{linear potential.} \tag{9.47}$$

m is the reduced mass. This equation can be checked with the hydrogen wave functions listed above. It would seem to be not that useful in the general case, since you have to know the wave functions and the potential to calculate either side. But the linear potential is special, since $dV/dr = \lambda$ is a constant. Substituting into Equation (9.30) we have for the linear potential:

$$\frac{\Gamma(J/\psi \to l^- l^+)}{\Gamma(\Upsilon(1S) \to l^- l^+)} = \frac{e_c^2}{e_b^2} \frac{m_c}{m_b} \frac{\lambda_c}{\lambda_b} \frac{M_\Upsilon^2}{M_{J/\psi}^2} = 7.7, \tag{9.48}$$

using the linear potential constants from Equation (9.42). The experimental ratio is 4.3, so our prediction is about a factor of two too large. This is an example of the application of Equation (9.47). For hydrogen-like wave functions — no confinement — we know that $|\psi_{10}(0)|^2 = 1/(\pi a_0^3)$. Equation (9.48) would then read:

$$\frac{\Gamma(J/\psi \to l^- l^+)}{\Gamma(\Upsilon(1S) \to l^- l^+)} = \frac{e_c^2}{e_b^2} \frac{M_\Upsilon^2}{M_{J/\psi}^2} \frac{(a_0(b\bar{b}))^3}{(a_0(c\bar{c}))^3} = 4.7.$$

This result is closer to the experimental number. The average radii of the $Q\bar{Q}$ atoms also agree well with their Bohr radii, based on the pure Coulomb-like potential in Equation (9.40). If a pure Coulomb-like potential is dominant at short distances, it is reasonable that the wave function at the origin is also Coulomb-like.

We have not succeeded in solving the complete problem of the quarkonium energy levels. The understanding of the $q\bar{q}$ force is of fundamental importance. The input data are limited, and working backwards in the Schrödinger equation is not easy. If you know the potential, you can in principle calculate the energy levels, and the wave functions. But the inverse problem of obtaining the potential given a few energy levels cannot be uniquely solved even in principle. So, we have to be satisfied with a discussion of the consequences of various hypotheses that can be treated at an introductory level, and hope that this encourages the interested reader to pursue the subject further.

9.5. K^0 Mixing

Mixing of the $K^0, \bar{K}^0$ neutral mesons was introduced in Section 1.9 of Chapter 1. The strong hypercharge $Y = N + S$ is conserved in production, so that K^0 with $Y = +1$, and the $\bar{K}^0$ with $Y = -1$ are produced in pairs, or a single K^0 is produced with a Λ hyperon. The cross-section for $\bar{K}$ in matter is larger than the cross-section for K, because the $\bar{K}$ can produce a hyperon $\Lambda, \Sigma^\pm$, or Σ^0. K mesons decay by weak interaction, and that breaks the $\Delta\, Y = 0$ conservation rule. K^0 and $\bar{K}^0$ have the same mass, but the two states with definite lifetimes can have different masses from the breaking of Y symmetry. The mixing can be described by a quantum mechanical mixing matrix, a form of the Schrödinger equation that incorporates the decay probability. We may write:

$$\left|K^0(t)\right\rangle = \begin{pmatrix} a(t) \\ 0 \end{pmatrix}, \quad \left|\bar{K}^0(t)\right\rangle = \begin{pmatrix} 0 \\ \bar{a}(t) \end{pmatrix},$$

$$-\frac{d}{dt}\begin{pmatrix} a(t) \\ \bar{a}(t) \end{pmatrix} = \begin{pmatrix} iM_{11} + \Gamma_{11}/2 & iM_{12} + \Gamma_{12}/2 \\ iM_{21} + \Gamma_{21}/2 & iM_{22} + \Gamma_{22}/2 \end{pmatrix}\begin{pmatrix} a(t) \\ \bar{a}(t) \end{pmatrix}. \tag{9.49}$$

Here, both M and Γ are Hermitian matrices, so $M_{21} = M_{12}^*$, and $\Gamma_{21} = \Gamma_{12}^*$, but the combination $iM + \Gamma/2$ is not Hermitian. The factor i appears on M rather than Γ because it has been moved from the time derivative. Both matrices have units of energy, but they are not even close to the same magnitude. Any particle that decays by weak interaction has a width that is much smaller than the particle mass. For example $m_{K^0} = 498\,\mathrm{MeV}$, and $\Gamma_{K_S} = \hbar/\tau = 7.3\times10^{-6}\,\mathrm{eV}$, more than 12 orders of magnitude smaller.

The simultaneous application in any order of the operations (Parity × Charge conjugation × Time reversal) leaves any sensible quantity in quantum field theory invariant. That is called the TCP theorem, and it guarantees that the masses and lifetimes of particles and antiparticles are the same, so:

$$M_{11} = M_{22}, \quad \text{and} \quad \Gamma_{11} = \Gamma_{22}. \tag{9.50}$$

If M_{12} and Γ_{12} are both real, then $iM + \Gamma/2$ can be easily diagonalized. Define eigenvalues $\lambda = \lambda_1 + i\lambda_2$ and solve two determinants with real eigenvalues. This works because the same unitary transformation diagonalizes each matrix. Then combine to obtain $\lambda_\pm$ and its eigenvectors:

$$\lambda_\pm = i(M_{11} \pm M_{12}) + (\Gamma_{11} \pm \Gamma_{12})/2, \quad \text{for } \lambda_+ : \ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\text{and for } \lambda_- : \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{9.51}$$

Equations (1.35) and (1.36) defined K_0 and $\bar{K}^0$ as linear combinations of the short and long lived neutral K mesons:

$$\begin{aligned} \left|K^0\right\rangle &= (|K_S\rangle + |K_L\rangle)/\sqrt{2}, \quad \text{and} \\ \left|\bar{K}^0\right\rangle &= (|K_S\rangle - \left|\bar{K}_L\right\rangle)/\sqrt{2}. \end{aligned} \tag{9.52}$$

Inverted we have:

$$\begin{aligned} |K_S\rangle &= (\left|K^0\right\rangle + \left|\bar{K}^0\right\rangle)/\sqrt{2}, \quad \text{and} \\ |K_L\rangle &= (\left|K^0\right\rangle - \left|\bar{K}^0\right\rangle)/\sqrt{2}. \end{aligned} \tag{9.53}$$

Comparing Equations (9.53) and (9.51) shows that λ_+ belongs to $|K_S\rangle$, and λ_- belongs to $|K_L\rangle$. The mass difference from the PDG is

$$\Delta m = m_L - m_S = -2M_{12} = 3.48 \times 10^{-6}\,\text{eV}. \tag{9.54}$$

That is a small number, but Γ is the same order of magnitude. We have for the decay rate eigenvalues:

$$\begin{aligned} \Gamma_S &= \Gamma_{11} + \Gamma_{12} = 7.35 \times 10^{-6}\,\text{eV}, \quad \text{and} \\ \Gamma_L &= \Gamma_{11} - \Gamma_{12} = 1.3 \times 10^{-8}\,\text{eV}. \end{aligned} \tag{9.55}$$

Obviously Γ_{11} and Γ_{12} are very nearly the same size: $\Gamma_{11} \sim \Gamma_{12} \sim \Gamma_S/2$, with Γ_{11} about 1/600 larger. This near equality of the diagonal and off-diagonal elements of Γ is due to the fact that the principal decay mode $K_S \to \pi + \pi$ directly couples $|K\rangle^0$ and $\left|\bar{K}^0\right\rangle$. When we begin discussion of mixing with the heavier quarks, c and b, there are hundreds of decay modes, and the fraction that are common to B^0 and $\bar{B}^0$ is very small, resulting in a much smaller difference between the short and long lifetimes.

Mixing leads to many different interesting phenomena. We reproduce Equations (1.37) and (1.38) here for convenience. The time development of the $\left|K^0(t)\right\rangle$ is given by

$$\left|K^0(t)\right\rangle = (e^{-t/(2\tau_s)} \times \left|K_S^0\right\rangle + e^{-t/(2\tau_L)} \times \left|K_L^0\right\rangle e^{-i\Delta m t})/\sqrt{2}. \tag{9.56}$$

At $t = 0$, $\left|K^0\right\rangle = \left|K^0\right\rangle$, and the probability of finding a $\left|\bar{K}^0\right\rangle$ at time t in a beam of pure $\left|K^0\right\rangle$ at $t = 0$ is

$$|\left\langle \bar{K_0} \middle| K_0 \right\rangle|^2 = (e^{-t/\tau_s} + e^{-t/\tau_L} - 2\cos(\Delta m t)e^{-(1/2\tau_s + 1/2\tau_L)t})/4. \tag{9.57}$$

The lifetime product $c\tau_L = 15.3$ m, and $c\tau_S = 2.7$ cm, so for distances of the order of one meter from the point of production it is safe to set $\tau_L \to \infty$. Then if we measure time in the K^0 rest frame in units of $\tau_S = 0.09$ nsec we have:

$$\begin{aligned} |\left\langle \bar{K_0} \middle| K_0 \right\rangle|^2 &= (e^{-t'} + 1 - 2\cos\left(\Delta m t'/\Gamma_s\right)e^{-t'/2})/4, \quad \text{and} \\ \Delta m/\Gamma_S &= 0.47. \end{aligned} \tag{9.58}$$

The cosine term vanishes at $t' = \pi\Gamma_s/(2\Delta m) = 3.3$, or about 0.3 ns. For a 1 GeV K^0 beam, $v\gamma c = 5.2 \times 10^{10}$ cm/s, so the K^0 in the laboratory has moved 15.6 cm from its production point when the $\cos(\Delta m t)$ term vanishes. Equation (9.58) is proportional to the probability of producing Λ and Σ hyperons if the K^0 beam is traveling through matter, a large bubble chamber for instance. Figure 9.9 shows this function vs distance from the production vertex in a bubble chamber for 1 GeV K^0. The distances are easy to measure, and the length is not that long to observe the bump from Δm oscillation. An experiment of this type measures $|\Delta m|$, but not its sign.

Another interesting phenomenon is regeneration. Regeneration is usually studied in a K_L^0 beam, that is far enough away from the production target to kill off all of the K_S^0, several meters, depending on the beam energy. Regeneration can also be studied in a beam of both K_L and K_S, with interference between direct $K_S \to \pi^+\pi^-$ and regenerated $K_S \to \pi^+\pi^-$, or in a gap configuration with the regenerator in two pieces, for interference between two regenerated beams. Last but not least regeneration has been used to study interference between the CP violating decay $K_L \to \pi^+\pi^-$ and regenerated K_S in the same beam. For a description of a regeneration experiment at Brookhaven National Laboratory, see [8] The total cross-section at modest energies for $\bar{K}^0$ is about 10 mb larger than it is for K^0 because of the inelastic channels producing hyperons. At lower energies the

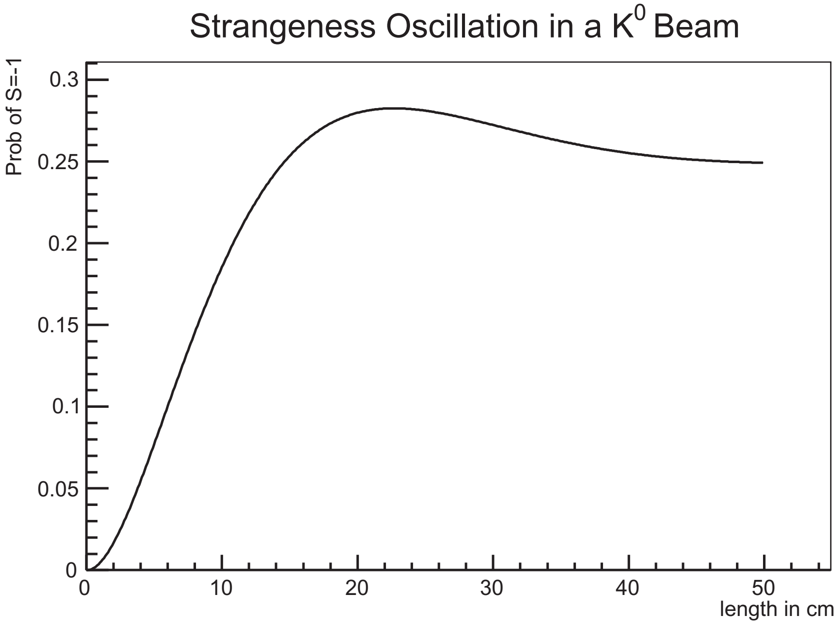


Figure 9.9: Equation (9.58), the probability of finding a $\bar{K}^0$ in a beam of 1 GeV K^0's as a function of distance from the production vertex. Half of the K^0's are K_L, and half of the K_L are $\bar{K}^0$, so the asymptotic value is 1/4. The bump is caused by the mass difference. The slow decrease due to the K_L lifetime is neglected.

K^-p cross-section has structure from hyperon resonances. At energies above 3 GeV the cross-sections are rather flat, but K^-p exceeds K^+p by about 50%. The optical theorem relates the total cross-section to the imaginary part of the forward scattering amplitude:

$$\sigma_{\text{tot}} = \frac{4\pi}{k}\text{Im}(f(0)), \quad \bar{\sigma}_{\text{tot}} = \frac{4\pi}{k}\text{Im}(\bar{f}(0)). \tag{9.59}$$

The wave number $k = p_K/\hbar$. The neutral K_L beam traveling through a solid target (the 'regenerator') coherently forward scatters off all of the target nuclei. If we write the mixed beam in the regenerator as

$$|\psi\rangle = a\,|K_L\rangle + b\,|K_S\rangle\,, \quad \text{then} \quad \rho = \frac{b}{a} = |\rho|e^{i\phi_\rho}.$$

At the output face of the regenerator

$$\rho = \frac{i\pi N\hbar^2 c^2(f(0) - \bar{f}(0))}{-m_K c^2(i\Delta mc^2 - \hbar/(2\tau_s))}(1 - e^{(i\Delta mc^2 - \hbar/(2\tau_s))L/(\hbar v\gamma c)}). \tag{9.60}$$

We have inserted normal units with $\hbar$ and c to facilitate numerical calculation. The difference in forward scattering amplitudes (not just the imaginary part) appears in the formula, but it is useful to estimate the magnitude of the effect by replacing $f(0) \to i\mathrm{Im}(f(0))$, because the total cross-sections are known. The real part of the amplitude should be small, and the regeneration phase can be measured. N is the number of nuclei per unit volume in the regenerator, and L is its length. The scattering amplitude has units of length. The ratio $L/v\gamma c$ is the time spent in the regenerator in the K^0 rest frame, and m_K is the K^0 mass. The mass difference $\Delta m = m_L - m_S$. The damping factor $e^{-t/(2\tau_S)}$ accounts for the fact that the K_S are decaying as they are produced in the regenerator. An equilibrium is reached in a thick regenerator, and the exponential disappears.

The Brookhaven experiment of Carithers *et al.*, noted above, operated in a K_L beam of average momentum $\langle p_K\rangle = 7.5\,\mathrm{GeV/c}$, or $v\gamma = 15$. The regenerator was made of carbon, 81 cm long, average density $1.72\,\mathrm{gm/cm^3}$. The mean decay length for 7.5 GeV/c K_S is about 40 cm, so for an 81 cm long regenerator $e^{-L/(v\gamma c\tau_S)} \sim 1/e$, but in our estimate we will ignore that term anyway. Using the total cross-sections for the forward scattering amplitudes, and dropping the exponential term gives:

$$|\rho| = \frac{N\hbar c(\sigma_T - \bar{\sigma}_T)v\gamma}{4\sqrt{(\Delta mc^2)^2 + \hbar^2/(2\tau_S)^2}}. \tag{9.61}$$

The numerator is

$$8.6 \times 10^{22}\ \mathrm{nuclei/cm^3} \times 1.9 \times 10^{-5}(\hbar c\ \mathrm{eV\text{-}cm}) \times 10^{-26}\ \mathrm{cm^2}\Delta\sigma_{\mathrm{tot}} \times 15(v\gamma) = 2.4 \times 10^{-7}\ \mathrm{eV},$$

and the denominator is

$$4\sqrt{(\Delta mc^2)^2 + (\hbar/(2\tau_S))^2} = 2 \times 10^{-5}\ \mathrm{eV},$$

giving

$$|\rho| = 0.012. \tag{9.62}$$

This is an approximate calculation, but it is safe to say that, even with a thick regenerator (almost 1 m of carbon, the order of one interaction length), the regenerated K_S amplitude is only a few percent, so the K_L beam is not severely attenuated. We will return to this phenomenon in the discussion of CP violation.

9.6. CP Violation for $K^0, \bar{K}^0$ Mesons

Speculation regarding colleagues in a remote corner of the universe who are doing sophisticated experiments in particle physics is amusing, and for the most part harmless. The direction of a true vector, like velocity, and the flow of time (breakfast, lunch, dinner) are, most people would agree, definitions that are unambiguous, even if our colleagues are composed of antimatter. Angular momentum poses a problem. It is a sense of rotation, and whether the vector is up or down relative to the plane of rotation is a matter of choice. Angular momentum, spin, and magnetic fields are all axial vectors, with directions specified by convention. Our colleague might well choose left handed rotation to be positive. It is a free choice. The direction of positive electric current is arbitrary. It was fixed before the discovery of the electron, so the electron had to have negative charge. What is the charge of a helium nucleus? Our colleagues could call it plus, even though we would say it is anti-helium. CP violation in the $K^0\bar{K}^0$ system is a small effect, requiring careful measurements in K^0 beams at particle accelerators like Brookhaven. Anti-Brookhaven could do the same experiments. We shall see that CP violation completely unravels the matter–antimatter puzzle, and that it is possible for Anti-Brookhaven to determine whether its nuclei are positively or negatively charged according to our convention. There is no experimental evidence for CP violation in the lepton sector. It has only been observed in weak decays of the quarks. However, once the K^0 is unquestionably identified, then because of strangeness conservation (or whatever our colleague calls it) the K^+ and K^- are also fixed, and the electron has the same charge as the K^-. So, our entire scheme of elementary particles can be sent to a remote station, and we may determine whether or not they are composed of antimatter.

Parity P and charge conjugation C are quantum mechanical observables with eigenvalues ±1. Parity is conserved by the strong and electromagnetic interactions, and was believed to be universally respected until the mid 1950s. Parity violation by the weak interactions was not widely anticipated before its discovery. But we are very fortunate that parity violation is a real, and in some sense maximal, effect, that leads to many interesting polarization phenomena. In this text, we have attempted to present the techniques necessary to calculate polarizations of spin 1/2 and spin 1 particles. Electrons are left-handed, and positrons are right-handed in weak interactions. Parity applied to the left-handed electron gives a right-handed electron that is forbidden by the weak interaction. However, if charge conjugation is also applied, one gets a right-handed positron, and that *is* allowed, so although we have lost P and C separately, we still have the product CP. If the TCP theorem holds, then CP is T. Time reversal is associated with a phase, as can be demonstrated with the non-relativistic Schrödinger equation, that is linear in time:

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi, \tag{9.63}$$

where H is a differential operator that is self-adjoint, satisfying

$$\int d^3x\phi^* H\psi = \int d^3x\psi H^*\phi^*. \tag{9.64}$$

The self-adjoint operation interchanges the functions, and takes the complex conjugate of the operator. For the Hamiltonian differential operator this condition fails if the potential is complex, as is explicitly true for our mixing matrix Equation (9.49). For a real potential the self-adjoint requirement gives a Hermitian matrix in the matrix representation. The self-adjoint Hamiltonian guarantees that the total probability is time independent.

$$\frac{\partial}{\partial t}\int d^3x\psi^*\psi = \frac{1}{i\hbar}\int d^3x(\psi^* H\psi - \psi H^*\psi^*) = 0. \tag{9.65}$$

If H is real in Equation (9.63), then reversing time and taking the complex conjugate gives the same equation for ψ^*:

$$i\hbar\frac{\partial\psi^*}{\partial t} = H\psi^*. \tag{9.66}$$

The only difference between ψ^* and ψ is a phase. Bound state wave functions that describe stable time independent states, in spherical coordinates have the form $\psi(r,\theta,\phi) = R(r)Y_l^{m_l}(\theta,\phi)$. The radial wave function can be chosen real. Time reversal flips the magnetic quantum number of the spherical harmonic — that is, it reverses the sense of rotation. Time reversal for a Dirac spinor was discussed in Chapter 2, Section 2.8. The explicit form of the matrix is Equation (2.180).

We return to our mass matrix, Equation (9.49), and diagonalize it again, but this time with one small complex phase on M_{12} of the mass mixing matrix. Introduce the two complex amplitudes p and q, thus,

$$\begin{aligned}|K_S\rangle &= \frac{1}{\sqrt{|p|^2+|q|^2}}(p\left|K^0\right\rangle + q\left|\bar{K}^0\right\rangle)\\ &= \frac{1}{\sqrt{2(1+|\epsilon|^2)}}((1+\epsilon)\left|K\right\rangle + (1-\epsilon)\left|\bar{K}\right\rangle),\end{aligned}$$

and

$$\begin{aligned}|K_L\rangle &= \frac{1}{\sqrt{|p|^2+|q|^2}}(p\left|K^0\right\rangle - q\left|\bar{K}^0\right\rangle)\\ &= \frac{1}{\sqrt{2(1+|\epsilon|^2)}}((1+\epsilon)\left|K\right\rangle - (1-\epsilon)\left|\bar{K}\right\rangle).\end{aligned} \tag{9.67}$$

These equations show two different ways to write the same thing. They are related by $\epsilon = (p-q)/(p+q) \ll 1$. The product $\langle K_S|K_L\rangle = (|p|^2-|q|^2)/(|p|^2+|q|^2) = 2\mathrm{Re}(\epsilon)/(1+|\epsilon|^2)$. If CP is conserved, p and q can both be chosen real, and $|p| = |q| = 1$, so $\langle K_S|K_L\rangle = 0$, and the two states are orthogonal. But if p and q are different complex numbers orthogonality is no longer guaranteed to hold. The notation p and q was introduced long ago by Lee, Oehme, and Yang [9], before even parity violation was discovered by experiment, much less

CP violation, but the seminal paper set the parameters for discussion of symmetry violations in weak interactions. They wrote down the following equations:

$$(iM + \Gamma/2)\psi_\pm = \lambda_\pm \psi_\pm, \quad \psi_\pm = \frac{1}{\sqrt{|p|^2 + |q|^2}} \begin{pmatrix} p \\ \pm q \end{pmatrix},$$
$$\lambda_\pm = iM_{11} + \Gamma_{11}/2 \pm (pq),$$

then

$$p^2 = \frac{\Gamma_{12}}{2} + iM_{12}, \quad \text{and} \quad q^2 = \frac{\Gamma_{12}^*}{2} + iM_{12}^*. \tag{9.68}$$

Equation (9.56) has to be modified in the presence of CP violation to preserve the normalization. For $\epsilon \neq 0$ we have:

$$\left|K^0(t)\right\rangle = \frac{\sqrt{2(1+|\epsilon|^2)}}{2(1+\epsilon)}\{|K_S\rangle\, e^{-t/(2\tau_S)} + |K_L\rangle\, e^{-t/(2\tau_L)} e^{-i\Delta m t}\}. \tag{9.69}$$

In our previous analysis, Γ_{12} and M_{12} were real, so $p^2 = q^2$. In order to introduce CP violation in the mass mixing matrix, we insert one very small phase, call it β, into the off diagonal mass mixing element M_{12}. To simplify the algebra, call $|M_{12}|e^{i\beta} = M(1 + i\beta)$, and $\Gamma_{12}/2 = \Gamma$. Then we have:

$$p^2 = \Gamma + iM - M\beta, \quad q^2 = \Gamma + iM + M\beta,$$
$$p = x + iy, \quad q = x' + iy'. \tag{9.70}$$

In order to solve for x and y we must confront quadratic equations in x^2 and y^2, so we have to take a lot of square roots, but bear with us. The reward is a formula for ϵ.

$$x^2 - y^2 = \Gamma - M\beta, \quad 2xy = M, \quad x^4 - (\Gamma - M\beta)x^2 - M^2/4 = 0,$$
$$x^2 = \frac{1}{2}(\Gamma + \sqrt{\Gamma^2 + M^2})\left(1 - \frac{M}{\sqrt{\Gamma^2 + M^2}}\beta\right),$$
$$y^2 = \frac{1}{2}(-\Gamma + \sqrt{\Gamma^2 + M^2})\left(1 + \frac{M}{\sqrt{\Gamma^2 + M^2}}\beta\right). \tag{9.71}$$

One more square root, treating $\beta \ll 1$:

$$x = \left\{(\Gamma + \sqrt{\Gamma^2 + M^2})/2\right\}^{1/2} \left(1 - \frac{M}{2\sqrt{\Gamma^2 + M^2}}\beta\right),$$
$$y = \left\{(-\Gamma + \sqrt{\Gamma^2 + M^2})/2\right\}^{1/2} \left(1 + \frac{M}{2\sqrt{\Gamma^2 + M^2}}\beta\right). \quad (9.72)$$

Now $x' = x(-\beta),\ y' = y(-\beta)$, so

$$p - q = \left\{-\left\{(\Gamma + \sqrt{\Gamma^2 + M^2})/2\right\}^{1/2} + i\left\{(-\Gamma + \sqrt{\Gamma^2 + M^2})/2\right\}^{1/2}\right\} \frac{M\beta}{\sqrt{\Gamma^2 + M^2}}, \quad (9.73)$$

and

$$p + q = 2\left\{\left\{(\Gamma + \sqrt{\Gamma^2 + M^2})/2\right\}^{1/2} + i\left\{(-\Gamma + \sqrt{\Gamma^2 + M^2})/2\right\}^{1/2}\right\}. \quad (9.74)$$

Finally we have:

$$\epsilon = \frac{p - q}{p + q} = \frac{M\beta}{2(\Gamma^2 + M^2)}(-\Gamma + iM)$$
$$= \frac{\mathrm{Im}(M_{12})}{2(\Gamma_{12}^2/4 + M_{12}^2)}(-\Gamma_{12}/2 + iM_{12}). \quad (9.75)$$

We have reinserted the mixing matrix element notation. Now $\Gamma_{K_S}/2 = \Gamma_{12} = 3.6 \times 10^{-6}\,\mathrm{eV}$, and the mass difference $\Delta m = -2M_{12} = 3.48 \times 10^{-6}\,\mathrm{eV}$, so numerically $\Gamma_{12}/2 \sim -M_{12}$, so the real part and the imaginary part are equal, and the phase is 45°, or $\pi/4$. M_{12} is negative because $M_L > M_S$. Using this phase Equation (9.75) can be written as follows:

$$\epsilon = \frac{\mathrm{Im}(M_{12})e^{i\pi/4}}{\sqrt{2}\Delta m}. \quad (9.76)$$

From the PDG we have:

$$|\epsilon| = 2.23 \pm 0.01 \times 10^{-3}, \quad \text{and}$$
$$2\mathrm{Re}(\epsilon) = 3.34 \pm 0.007 \times 10^{-3}, \quad \text{or} \quad \phi_\epsilon \sim 42^\circ. \quad (9.77)$$

From Equation (9.67) we may calculate $\eta_\pm$, the parameter that characterizes the discovery of CP violation.

$$\left\langle \pi^+\pi^- \middle| K_L \right\rangle = \frac{\epsilon}{\sqrt{2}}(\left\langle \pi^+\pi^- \middle| K^0 \right\rangle + \left\langle \pi^+\pi^- \middle| \bar{K}^0 \right\rangle) = \sqrt{2}\epsilon(\left\langle \pi^+\pi^- \middle| K^0 \right\rangle,$$

$$\left\langle \pi^+\pi^- \middle| K_S \right\rangle = \sqrt{2}\left\langle \pi^+\pi^- \middle| K^0 \right\rangle,$$

$$\eta_\pm = \frac{\langle \pi^+\pi^- | K_L \rangle}{\langle \pi^+\pi^- | K_S \rangle} = \epsilon. \tag{9.78}$$

There is a small correction to this ratio called ϵ' that accounts for what is called ‘direct’ CP violation from interference between two I-spin amplitudes for $\pi^+\pi^-$. This quantity exists, but is only about $10^{-3} \times \epsilon$, and we may neglect it in a simplified treatment. Most of the CP violation in the neutral K sector comes from the mass mixing matrix. The discovery experiment reported $(K_L \to \pi^+\pi^-)/(K_L \to$ all charged modes$) = 2.0 \pm 0.4 \times 10^{-3}$. This ratio is close to $\eta_\pm$, but it is *not* $\eta_\pm$. Numerical cancellations make the measured ratio almost equal to the amplitude ratio:

$$\begin{aligned}\Gamma(K_L \to \pi^+\pi^-) &= \Gamma(K_S \to \pi^+\pi^-) \times |\eta_\pm|^2 \\ &= 0.686 \times \Gamma(K_S) \times 4.97 \times 10^{-6} \\ &= 2.5 \times 10^{-11}\,\text{eV}, \\ \Gamma(K_L \to \text{all charged modes}) &= 0.8 \times \Gamma(K_L) = 1.04 \times 10^{-8}\,\text{eV}.\end{aligned} \tag{9.79}$$

The ratio is 2.4×10^{-3}, almost exactly equal to $\eta_\pm$! The numerical factors 0.686 and 0.8 are branching ratios for $K_S \to \pi^+\pi^-$ and $K_L \to$ (all charged modes), respectively. The dominant neutral modes are $K_S \to \pi^0\pi^0 = 31\%$ and $K_L \to \pi^0\pi^0\pi^0 = 20\%$.

$K^0 \to \pi^+\pi^-$ has special characteristics that enable it to be identified in a neutral K_L^0 beam in the presence of backgrounds several hundred times larger. The two charged pion momenta can be measured with tracking chambers and a magnet. The two input tracks should point back to a vertex somewhere along the neutral beam. The two pion momentum vectors should form a plane that contains the neutral beam, and sum to a vector that lies in the neutral beam, and points back to the production target. Assuming that the particles are pions, one can calculate the invariant mass squared: $s = (E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2$, and the result should be close to $m^2_{K^0}$. For the CP violating decay $K^0_L \to \pi^+\pi^-$ the backgrounds are three-body decays $\pi\mu\nu, \pi e\nu$, and $\pi^+\pi^-\pi^0$. These background rates are much larger than the $\pi^+\pi^-$ signal, but the missing third particle carries off momentum. The semileptonic three-body decays have a missing neutrino. The two charged particles ($\pi^+\mu^-$ for example), form a vertex that is in the neutral beam, but because of the neutrino the total momentum vector $\vec{p}_\pi + \vec{p}_\mu$ does not lie along the beam line. If the beam direction is assumed, then there is a two-fold ambiguity regarding the parent K_L momentum that corresponds to forward or backward neutrinos in the K_L center of mass. It is the overconstrained kinematics of the two-body decay that facilitated the discovery of the CP violating $K^0_L \to \pi^+\pi^-$. To repeat the arguments from Chapter 1:

$$\begin{aligned} &\mathrm{CP}\left|K^0\right\rangle = \left|\bar{K}^0\right\rangle, \quad \text{if} \quad \epsilon = 0, \quad \mathrm{CP}\left|K_S\right\rangle = +\left|K_S\right\rangle, \\ &\mathrm{CP}\left|K_L\right\rangle = -\left|K_L\right\rangle, \quad \text{and} \quad \mathrm{CP}\left|\pi^+\pi^-\right\rangle = +\left|\pi^+\pi^-\right\rangle. \end{aligned} \tag{9.80}$$

Therefore, $K_S \to \pi^+\pi^-$ is allowed, but $K_L \to \pi^+\pi^-$ is forbidden. Figure 9.10 from the discovery paper of Christenson *et al.* [10] shows the angular distribution of events in three mass bins: below, on, and above the K^0 mass. θ is the angle with respect to the neutral beam direction, defined by collimators. $\cos\theta = 0.9999$ is about 0.8 degree.

The charge asymmetry in $K_L \to \pi\mu\nu$ or $K_L \to \pi e\nu$ can be derived directly from Equation (9.67), assuming the validity of the $\Delta Q = \Delta S$ rule, that is predicted by the SU3 octet operators. $\Delta Q = \Delta S$ predicts $\langle\pi^-| H_W \left|K^0\right\rangle$ and $\langle\pi^+| H_W \left|\bar{K}^0\right\rangle$ are non-vanishing, but the two matrix elements with the pion charge interchanged vanish. Look at the pseudoscalar meson octet to see the rule. The presence of $\Delta Q = -\Delta S$ would decrease the sensitivity of the charge asymmetry to $\mathrm{Re}(\epsilon)$, but we ignore such an effect because the experimental limits on the amplitude for $\Delta Q = -\Delta S$ in K^0 decay are less than 1%. If the $\Delta Q = \Delta S$ rule did not hold, but $\epsilon = 0$, there would be no charge asymmetry. From Equation (9.67):

$$\left\langle\pi^+\mu^-\bar{\nu}_\mu\right| H_W \left|K_L\right\rangle = -\frac{1}{\sqrt{2}}(1-\epsilon)\left\langle\pi^+\mu^-\bar{\nu}_\mu\right| H_W \left|\bar{K}^0\right\rangle,$$

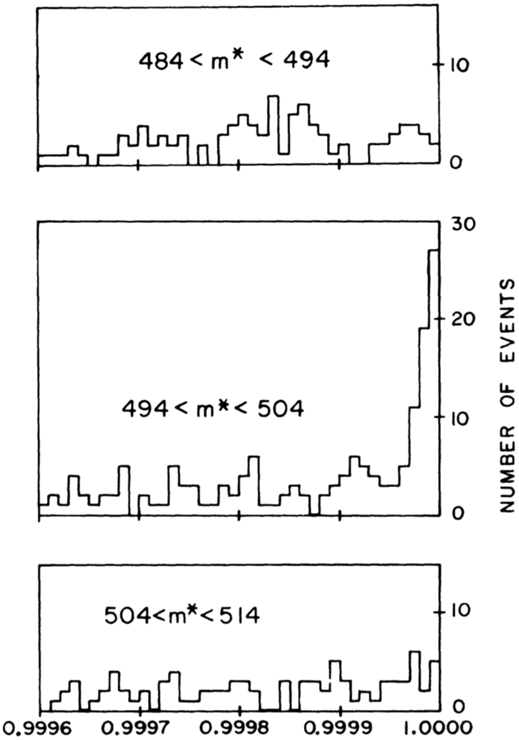


Figure 9.10: Original data for the discovery of CP violation in the decay $K_L \to \pi^+\pi^-$. The three plots are mass bins for the $\pi^+\pi^-$ pair. The K^0 mass is 497.6 MeV, contained in the center plot. Lower mass events come from dominant three body decays $K_L \to \pi^+\pi^-\pi^0$. Higher masses are mis-measurements. The x axis is $\cos\theta$, where θ is the polar angle between the vector sum $\vec{p}_+ + \vec{p}_-$ and the direction of the neutral beam, called 'target pointing'. The K_L momentum points back to the production target. The signal is the peak near $\cos\theta = 1$ in the center plot. Credit Ref. [10].

and

$$\left\langle \pi^-\mu^+\nu_\mu \right| H_W \left| K_L \right\rangle = \frac{1}{\sqrt{2}}(1+\epsilon) \left\langle \pi^-\mu^+\nu_\mu \right| H_W \left| K^0 \right\rangle . \tag{9.81}$$

The charge asymmetry for the lepton is defined by

$$A_\pm = \frac{|\left\langle \pi^-\mu^+\nu_\mu \right| H_W \left| K_L \right\rangle|^2 - |\left\langle \pi^+\mu^-\bar{\nu}_\mu \right| H_W \left| K_L \right\rangle|^2}{\text{sum}} = 2\mathrm{Re}(\epsilon), \tag{9.82}$$

where we have ignored $|\epsilon|^2 = 4 \times 10^{-6}$ in the denominators. The experimental number is given in Equation (9.77) above. This CP violation effect alone is sufficient to determine the signs of all of the charges in an independent and unambiguous way, so that whether a distant laboratory is matter or antimatter can be established. There is, however, another effect that does the same thing, and that is interference between regeneration and CP violation.

There are two papers by Carithers *et al.*, referred to above. The first describes the regeneration, and the measurement of the regeneration phase. The second paper presents the results of a study of the interference between the regenerated $K_S \to \pi^+\pi^-$ and the CP violating decay $K_L \to \pi^+\pi^-$. The formula for the signal as a function of time measured from the output face of the regenerator in the K^0 rest frame with a suitable overall normalization constant C is

$$\begin{aligned} I(K \to \pi^+\pi^-) = C(&|\rho|^2 e^{-\Gamma_S \tau} + |\eta_\pm|^2 e^{-\Gamma_L \tau} \\ &+ 2|\rho||\eta_\pm| e^{-(\Gamma_S+\Gamma_L)\tau/2} \cos(\Delta m \tau + \phi_\rho - \phi_\pm)). \end{aligned} \tag{9.83}$$

The formula for ρ is Equation (9.60). The phases are $\phi_\rho = -40.9 \pm 2.6^\circ$, and $\phi_\pm = 45^\circ$, so $\phi_\rho - \phi_\pm \sim -\pi/2$, and the cosine term becomes $\sin(\Delta m \tau)$. The mass difference $\Delta m = 30 \times 10^{10}$ degrees/s, so $\Delta m \tau = 180^\circ$ at 6×10^{-10} s. The pronounced drop at about 6.5 K_S lifetimes in Figure 9.11 is from $|\rho||\eta_\pm|$ interference, which is negative in this region. It is this effect that establishes which particle is K^0 and which is $\bar{K}^0$ on a remote planet. Suppose the experiment is setup at anti-Brookhaven. Then the regenerator is anti-carbon, and since the strong interaction conserves CP, the cross-sections $K^0 + p$ and $\bar{K}^0 + \bar{p}$ are the same. Then in the expression for ρ, $f(0) - \bar{f}(0) \to \bar{f}(0) - f(0)$, or the amplitude is multiplied by $e^{i\pi} = -1$. This sign change turns the drop into a bump. Problem 9.8 is a numerical exercise demonstrating this effect. Thus can the ambiguity between particle and antiparticle be resolved. If the interference between the regenerated and CP violating amplitudes is negative, the regenerator is ordinary matter; if it is positive, the regenerator is antimatter. It is that simple.

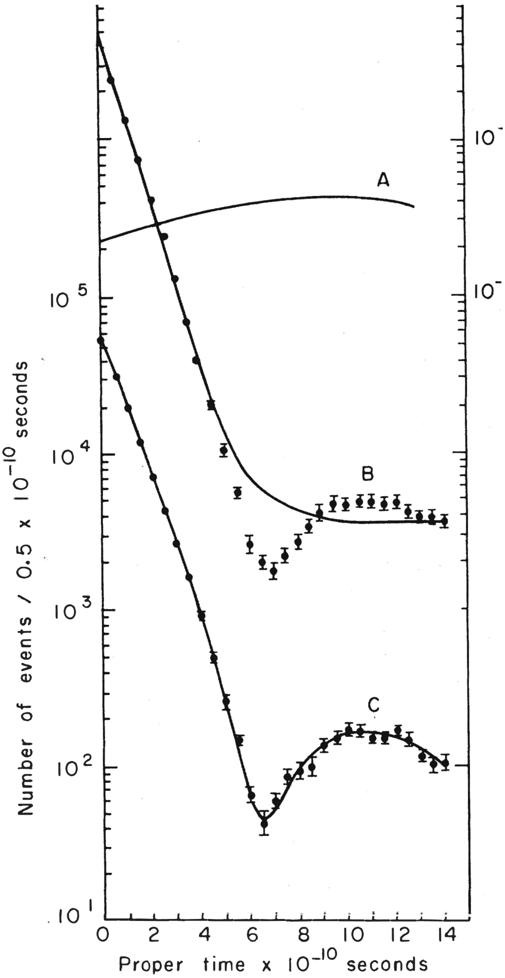


Figure 9.11: Results of the measurement of interference between the regeneration amplitude ρ and the CP violating amplitude $\eta_{\pm}$ in a K_L beam of average momentum 5.5 GeV/c at Brookhaven National Laboratory. The acceptance as a function of proper time measured from the output face of the regenerator is the top A curve, with numbers to the right (10^{-1} and 10^{-2}). The K_S lifetime is almost 10^{-10} s. When the regenerated K_S amplitude decays to the same size as the CP violating K_L amplitude (about six lifetimes) the two destructively interfere, giving the sharp drop in the $K \to \pi^+\pi^-$ signal, then the oscillation above the tail. At long times there is a flat tail from $\eta_{\pm}$ alone. The solid B curve shows the expected yield with no oscillation. The bottom curve is the same, without various corrections, to show that the interference was there in plain sight. C curve is Equation (9.83). Credit [7, p. 1244].

9.7. Mixing and CP Violation in $\Upsilon(4S) \to B\bar{B}$

The quantum number that acts like charge to distinguish K^0 and $\bar{K}^0$ is called strangeness, S. Heavy flavor quarks c and b have analogous quantum numbers conserved by strong and electromagnetic interactions, but violated by the weak interactions, called charm C and bottom B, and heavy neutral mesons containing these quarks can mix in the same way as the $K\bar{K}$. Separate two by two mixing matrices describe each heavy $q\bar{q}$ pair. Therefore, $D^0\bar{D}^0$ and $B^0\bar{B}^0$ both mix. Although we will discuss in some detail only the $B_d = d\bar{b}$ mixing and CP violation, we have summarized for reference the mixing parameters of all of the heavy quark mesons in Table 9.5.

All of these numbers have been adopted from listings by the PDG. The first column lists only the particle. Mixing is $K\bar{K}$, etc. Only pairs of neutral particles mix. Short and long lived versions have not been observed for B_d, and are very close together for D^0. Uncertainties are mostly omitted, except where the number has large errors, like ΔM for the D^0. The B_s is an ordinary B^0 with the down quark replaced by a strange quark. The mass is slightly heavier, and $B_s\bar{B}_s$ is too heavy to be fed by the $\Upsilon(4S)$ at 10,579 MeV. The B_s has an unusually large mass difference, that leads to very rapid oscillations (25 in one mean life). Since the mean life is 1.5 ps, we are talking about a lot of wiggles in a very short time. For this reason among others most of the studies of B_s have been carried out at hadron colliders like the Tevatron and LHC, where there is some help from the Lorentz factor γ on the short times. Research is still in progress at the LHC, where a dedicated experiment LHCb is studying heavy

Table 9.5: Heavy quark mixing.

Meson	Flavor change	$\Gamma_S(10^{-6}\,\text{eV})$	$\Gamma_L\ (10^{-6}\,\text{eV})$	M(MeV)	$\Delta M\ (10^{-6}\,\text{eV})$	$\Delta M/\Gamma_S$
$K(d\bar{s})$	$\Delta S = 2$	7.35	1.28×10^{-2}	497.6	3.48	0.47
$D(c\bar{u})$	$\Delta C = 2$	1700	$\Delta\Gamma/\Gamma = 0.013$	1864.8	5.9 ± 2.6	0.003
$B_d(d\bar{b})$	$\Delta B = 2$	433	$\Gamma_L = \Gamma_S$	5279.6	333.	0.77
$B_s(s\bar{b})$	$\Delta B = 2$	463	406	5366.9	11,600	25

flavor decays. In the following discussion B^0 means B_d^0 unless otherwise noted.

Heavy flavor mixing was first discovered by UA1 at the CERN $p\bar{p}$ collider by comparing like sign and opposite sign semi-leptonic decays from pair production of heavy mesons. For a detailed discussion we have selected $\Upsilon(4S) \to B^0\bar{B}^0$ because the initial state has known quantum numbers $J^{PC} = 1^{--}$, and this leads to interesting consequences for the mixing.

Figure 9.5 shows the rather broad and not very pronounced $\Upsilon(4S)$ resonance at 10,579 MeV, with a width $\Gamma(4S) = 20\,\mathrm{MeV}$. At the $\Upsilon(4S)$ mass one unit of $R = (4\pi\alpha^2)/(3s) = 1.1 \times 10^{-33}\,\mathrm{cm}^2$ (1.1 nb). From the inset curve one can see that on the resonance $R = 6.2$ with a background level of 4.5, so sitting on the resonance gives a signal to total of about 30%. The background events are back-to-back dijets from $e^-e^+ \to q\bar{q}$ light quarks, that can be distinguished from the spherical track distribution of $B\bar{B}$ decay. The decay $\Upsilon(4S) \to B\bar{B}$ has a branching ratio greater than 96%, and half decay to $B^0\bar{B}^0$, the other half to B^+B^-. The total cross-section on the resonance is $6.2 \times 1.1 = 6.8 \times 10^{-33}\,\mathrm{cm}^2$, giving about $1 B^0\bar{B}^0$ pair per second at a luminosity of $10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The Belle collaboration [11] reports 31.3×10^6 $B^0\bar{B}^0$ pairs from an integrated luminosity $\int \mathcal{L} dt = 29 \times 10^{39}\,\mathrm{cm}^2$ (about a year of running), for an effective cross-section of $\sigma = 1\,\mathrm{nb}$, consistent with these numbers.

The $B\bar{B}$ pair is produced almost at rest in the center of mass of the $\Upsilon(4S)$. The velocity of either B is

$$\begin{aligned} &E_B = M_\Upsilon/2 = 5289.5\,\mathrm{MeV}, \\ &M_B = 5279.6\,\mathrm{MeV}, \quad p = 323\,\mathrm{MeV/c}, \quad v = 0.061. \end{aligned} \tag{9.84}$$

The B^0 lifetime is $\tau_B = 1.5 \times 10^{-12}\,\mathrm{s}$, so the mean decay distance is only $30\,\mu$m! A 5 GeV K_L has a mean decay length of 150 m, but B^0 mesons from $\Upsilon(4S)$ decay do not travel very far. If the spatial resolution of the vertex detector is about $30\,\mu$m, which is quite good, then both B and $\bar{B}$ often appear to decay at the same point in space, giving lots of tracks to sort through. Energy, momentum and charge balance are kinematic quantities that help assign tracks to either the B or the $\bar{B}$. The B mass is about ten times the K_S mass, and there are literally 100's of decay modes in the PDG listings, most of them with very small branching ratios. Dominant decay modes are $B^0 \to D^- + X$ and $B^0 \to \bar{D}^0 + X$, where $d\bar{b} \to d\bar{c}$ or $u\bar{c}$. Short distances, high multiplicity, and the need for particle identification (π, K, p) place stringent requirements on detectors built to study B mesons at the $\Upsilon(4S)$.

There are several different decay modes of the B^0 that can serve as a 'flavor tag', that is the final state determines whether the parent was a B^0 or a $\bar{B}^0$. A charmed meson in the final state is a good flavor tag, because $\bar{b} \to \bar{c}$ and $b \to c$ means $B^0 \to \bar{D}^0$ or D^-, and charge conjugates for $\bar{B}^0$. Charged leptons are another flavor tag: $B^0 \to l^+$ and $\bar{B}^0 \to l^-$, because $\bar{b} \to \bar{c} + W^+$. The lepton charge tag was used in the UA1 study of heavy quark mixing referred to above. There are some rare final states that are common to both B^0 and $\bar{B}^0$. One that is simple to understand on the quark level is

$$B^0 \to J/\psi K_S \to \bar{B}^0, \quad \text{at the quark level } d\bar{b} \to c\bar{c} + d\bar{s},$$

and

$$b\bar{d} \to c\bar{c} + s\bar{d}. \tag{9.85}$$

The K_S is an equal mixture of $s\bar{d}$ and $d\bar{s}$. The final state $J/\psi K_S$ is an eigenstate of CP with CP $= -1$. The J/ψ is $J^{PC} = 1^{--}$, so it is CP even. Since the B^0 and K_S are both pseudoscalars, the final state has orbital angular momentum one, and odd parity. The K_S contains equal amounts of K^0 and $\bar{K}^0$ and is CP even. The K_L, also a possible final state but more difficult to detect, is CP odd, so that $J/\psi K_L$ is CP even. However, the branching ratios are only about 5×10^{-4}, so you need to make a lot of $\Upsilon(4S)$ to detect a sample of these common modes. The decay $J/\psi \to l^- l^+$ for detection of the J/ψ costs another factor of ten. The Belle experiment mentioned above had a sample of 457 fully reconstructed $B^0/\bar{B}^0 \to J/\psi K_S$, and 569 $J/\psi K_L$ from 3×10^7 $B\bar{B}$ pairs. This small branching ratio is to be distinguished from the K^0 system, where $K \to \pi\pi \to \bar{K}$ is dominant. Because of the small probability of final states that can mix, the matrix element Γ_{12} in Equation (9.49) is small, and the B_S and B_L lifetimes are almost the same. In fact that labelling is not used, rather they are B_H and B_L for 'heavy' and 'light'. The mass difference, that governs the oscillation frequency in the mixing, is

$$\Delta m_B = m_{B_H} - m_{B_L} = 3.3 \times 10^{-4}\,\text{eV}, \quad \frac{\Delta m_B}{\Gamma_{B^0}} = 0.769 \pm 0.004. \tag{9.86}$$

The ratio is comparable to $\Delta m/\Gamma_S$ for K mesons. The B meson mass difference is about 100 times larger that the K meson mass difference, and so is the total decay rate. Equations (9.56) and (9.57) for B^0 mixing are written as follows:

$$\left|B^0(t)\right\rangle = e^{-t/(2\tau_0)}(\left|B_L^0\right\rangle + \left|B_H^0\right\rangle e^{-i\Delta m t})/\sqrt{2}, \tag{9.87}$$

and

$$\left|B_L^0\right\rangle = (\left|B^0\right\rangle + \left|\bar{B}^0\right\rangle)/\sqrt{2}, \quad \left|B_H^0\right\rangle = (\left|B^0\right\rangle - \left|\bar{B}^0\right\rangle)/\sqrt{2},$$

$$|\left\langle \bar{B}^0 \middle| B^0 \right\rangle|^2 = \frac{e^{-t/\tau_0}}{2}(1 - \cos(\Delta m t)). \tag{9.88}$$

There is only one lifetime, and it is 1.5×10^{-12} s, so these oscillations are fast. We have already noted that the velocity of either B or $\bar{B}$ is very small from the decay of the $\Upsilon(4S)$, and it is more picturesque to imagine them flying apart to some great distance for what we are about to say. The point is that the quantum numbers of the $\Upsilon(4S)$ are 1^{--}, so the $B^0\bar{B}^0$ pair is odd under charge conjugation, and therefore cannot mix. It is impossible for two B's or two $\bar{B}$'s to be odd under charge conjugation. It is the decay of one member of the pair that releases the other from prison and allows it to start mixing. This is a perfect example of quantum entanglement. The pair is produced at $t = 0$, and the $\bar{B}^0$ decays into a tagged final state, $l^-\bar{\nu}_l + X$ for example, at $t = t_0$. At that time the other meson is definitely a B^0. Equation (9.89) then applies, with $\Delta t = t - t_0$ in the formula:

$$|\left\langle \bar{B}^0 \middle| B^0 \right\rangle|^2 = \frac{e^{-|\Delta t|/\tau_0}}{2}(1 - \cos(\Delta m \Delta t)). \tag{9.89}$$

As in all of these formulas time is measured in the B^0 rest frame. We will be interested below in an experiment where one B decays as a tag, and the other decays into a common final state that is accessible to both B and $\bar{B}$. In that case Δt can be either positive or negative, hence the magnitude sign in the exponent of Equation (9.89).

CP violation is usually a small phase, like β that was introduced into the off diagonal term in the 2×2 mass matrix for the $K\bar{K}$ system. That phase was inserted by hand — not predicted by some grand theory. For the heavy quarks there is another place where, again by hand, a phase may be inserted to cause CP violation. The strange quark is one of the light quarks, and it is difficult to link it to the K meson with the same confidence that the bottom quark can be linked to the B meson. A phase in the CKM matrix was introduced in Chapter 7, Section 7.5, with the following sentence: "CP violation in the weak interaction is described by a single complex phase in the CKM matrix." We shall now explore that assertion in more detail. One of its consequences is that CP violation is restricted to quarks, and there are no predictions regarding possible CP violation in the lepton sector — a open question for experimental study. Two forms of the CKM quark mixing matrix are given in Chapter 7, and we repeat them here for convenience:

The quark label form

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.97446 & 0.22452 & 0.00365 \\ 0.22438 & 0.97359 & 0.04214 \\ 0.0089 & 0.04133 & 0.999105 \end{pmatrix}. \tag{9.90}$$

The second matrix contains the magnitudes, without the complex phases. Also we have omitted the errors for simplicity. The errors are given in Chapter 7. The Wolfenstein form with the complex phase to order λ^3 is

$$V_{\mathrm{Wolf}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{9.91}$$

In the Wolfenstein form η is the imaginary part, that represents the CP violating phase. Note that it appears in opposite corners — V_{ub} and V_{td}, and so spans all three quark flavors. There is a reason for this. A 2 × 2 matrix, Cabibbo's hypothesis with one more charge +2/3 quark, is not large enough to have a free phase, but a 3 × 3 matrix is, so the phase has to range over all three quarks. The reason may seem obscure, but it is a basic property of unitary matrices and the arbitrariness of quark phases in the Lagrangian. A 2 × 2 unitary matrix with two arbitrary phases and S_{11} real has the form:

$$S = \begin{pmatrix} S_{11}e^{i\delta} & \sqrt{1-S_{11}^2}e^{i\beta} \\ \mp\sqrt{1-S_{11}^2}e^{-i\beta} & \pm S_{11}e^{-i\delta} \end{pmatrix}, \quad \text{and}$$

$$S^\dagger = \begin{pmatrix} S_{11}e^{-i\delta} & \mp\sqrt{1-S_{11}^2}e^{i\beta} \\ \sqrt{1-S_{11}^2}e^{-i\beta} & \pm S_{11}e^{i\delta} \end{pmatrix}. \tag{9.92}$$

The matrix multiplies d and s quarks, and the two phases β and δ may be absorbed in the quark spinors without any change, because in the Lagrangian the spinors appear as $\bar{u}u$. So the 2 × 2 matrix can be written in terms of a single real angle θ as

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{9.93}$$

The (d, s) quarks are the mass eigenstates, and (d', s') couple to the charge 2/3 quarks. The coupling to a W^- vector boson would then be of the form:

$$\begin{pmatrix} \bar{u} & \bar{c} \end{pmatrix} \frac{g}{\sqrt{2}}\gamma^\mu(1-\gamma^5)/2 \begin{pmatrix} d' \\ s' \end{pmatrix} \tag{9.94}$$

This is Cabibbo's idea with an extra charge +2/3 quark, charm, added to make the 2 × 2 matrix unitary. The unitary matrix forbids flavor changing neutral currents among the quarks. Thus,

$$\begin{pmatrix} \bar{d}' & \bar{s}' \end{pmatrix} = \begin{pmatrix} \bar{d} & \bar{s} \end{pmatrix} S^\dagger, \quad \begin{pmatrix} \bar{d}' & \bar{s}' \end{pmatrix} \begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \bar{d} & \bar{s} \end{pmatrix} S^\dagger S \begin{pmatrix} d \\ s \end{pmatrix} = \bar{d}d + \bar{s}s. \tag{9.95}$$

This is an important property of a unitary quark mixing matrix for the weak interaction, because flavor changing neutral currents like $K^+ \to \pi^+\nu\bar{\nu}$, which could be comparable to the dominant mode $K^+ \to \pi^0\mu^+\nu_\mu$, are strongly suppressed.

The paper of Kobayashi and Maskawa is dated before the discovery of either the charm or the bottom quark, so they were very prescient. They realized that with six quarks (we were half way there then, with (u, d, s)), there would be one free phase. A clear argument as to why this is so was given by Toshihide Maskawa in his Nobel Lecture (2008). Suppose we change the phases of the quarks:

$$\begin{aligned} &u_i \to e^{ia_i} u_i, \quad d_j \to e^{ib_j} d_j, \quad \text{then the matrix element} \\ &V_{ij} \to e^{-i(a_i - b_j)} V_{ij}. \end{aligned} \tag{9.96}$$

If there are $2N$ quarks, there are $2N$ phase changes, but a common phase change does not affect the matrix V_{ij}, so there are $2N - 1$ degrees of freedom to change the phases on V. An $N \times N$ unitary matrix has $N(N + 1)/2$ degrees of freedom, and subtracting $2N - 1$ gives:

$$\frac{N(N+1)}{2} - (2N-1) = \frac{(N-1)(N-2)}{2}. \tag{9.97}$$

This vanishes for $N = 1$ (two quarks), and for $N = 2$ (four quarks), but is equal to one for $N = 3$ (six quarks). Hence one needs at least six quarks to introduce CP violation in the CKM mixing matrix.

This discovery motivated an acute interest in the properties of 3×3 unitary matrices. Referring to Equation (9.90) and its Hermitian conjugate, we have three equations of products that equal one, and three independent equations that equal zero (the other three are complex conjugates):

$$\begin{pmatrix} V_{ud}^* & V_{cd}^* & V_{td}^* \\ V_{us}^* & V_{cs}^* & V_{ts}^* \\ V_{ub}^* & V_{cb}^* & V_{tb}^* \end{pmatrix} \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{9.98}$$

The three off diagonal products that add to zero are:

$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0, \tag{9.99}$$

$$V_{ud}^* V_{us} + V_{cd}^* V_{cs} + V_{td}^* V_{ts} = 0, \tag{9.100}$$

and

$$V_{us}^* V_{ub} + V_{cs}^* V_{cb} + V_{ts}^* V_{tb} = 0. \tag{9.101}$$

It is a property of the 3×3 unitary matrix that all three of these equations may be plotted in a two dimensional complex space to form closed triangles, and that the areas of these triangles are all the same. One phase angle opens up all of the triangles. If CP is conserved, all triangles are flat, with zero area. The top relation, Equation (9.99), is the one of special interest. It is the one that gives a triangle in the Wolfenstein approximation. A glance at the Wolfenstein form shows that to order λ^3 the elements V_{td} and V_{ub} are complex, and all of the other elements are real. Also V_{ud} and V_{tb} are very close to one, so to a good approximation this equation reads:

$$\begin{gathered} V_{ub}^* + V_{cb} V_{cd} + V_{td} = 0; \text{ or for Wolfenstein} \\ A\lambda^3(\rho + i\eta) - A\lambda^3 + A\lambda^3(1 - \rho - i\eta) = 0. \end{gathered} \tag{9.102}$$

The sum of the one real and two complex numbers does indeed vanish, as Wolfenstein's form is unitary to order λ^3, but we retain the three terms and plot them as a triangle in a two dimensional complex space. The area of this triangle is $(A\lambda^3)^2\eta/2 \sim 1.5 \times 10^{-5}$. The other two triangles, Equations (9.100) and (9.101), are flat in Wolfenstein's approximation, because the complex phases in V_{cs}, and V_{cb} are multiplied by λ^4. This does not violate the claim that all three triangles have the same area; Wolfenstein's approximation is unitary only to order λ^3. You might ask why it was that this particular form of approximation was chosen. The reason is that it introduces the CP violating phase at the lowest order, namely λ^3. Wolfenstein's original paper [12] gives an improved form where the imaginary part is expanded to order λ^5, and this fix inflates the other two triangles, so that the areas are all the same. The new triangles are right triangles to good approximation. This improvement is left as an exercise in Problem 9.13. The PDG gives the exact form for the matrix, but it is rarely used for discussion of CP violation. From the PDG listings the best fits to Wolfenstein's form, given in Chapter 7, are $A = 0.84, \lambda = 0.224, \rho = 0.14, \text{and } \eta = 0.36$, so the common multiplier $A\lambda^3 = 0.009$. The small multiplier $A\lambda^3$ means that this triangle

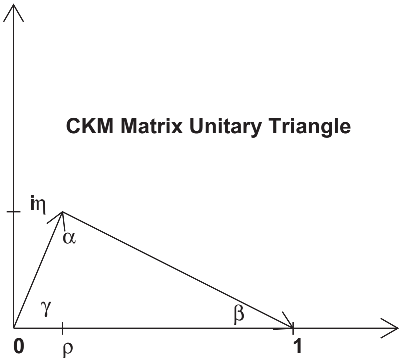


Figure 9.12: CKM 'golden' triangle with $A\lambda^3$ factored out. The angles α, β, γ are called ϕ_2, ϕ_1, ϕ_3 in the Belle papers. The angles on the drawing are: $\alpha = 88^\circ, \beta = 23^\circ$, and $\gamma = 69^\circ$, while the sides are: base = 1, left side = 0.36, and right side = 0.93. These lengths are related to magnitudes of CKM matrix element products $V_{cb}^* V_{cd}, V_{ub}^* V_{ud}$, and $V_{tb}^* V_{td}$, in the Wolfenstein approximation.

has large opening angles, to have the same area as other triangles with longer sides. It is for this reason that this triangle is singled out — it should show the imaginary phase at its best. It is called *the* 'unitary triangle', and in the Wolfenstein form with $A\lambda^3$ factored out is drawn in Figure 9.12. If the triangle sides are (a, b, c), then the base $a = 1$ is real, and the complex numbers have magnitudes $|b| = \sqrt{(1-\rho)^2 + \eta^2} = 0.93$; $|c| = \sqrt{\rho^2 + \eta^2} = 0.38$. This is the way it is drawn, but all of the numbers are experimental measurements, and the closing of the triangle has uncertainties. The angles are labeled α at the top, β between $1 - \rho - i\eta$ and one, and γ between $\rho + i\eta$ and one. In terms of the unitary matrix elements we have:

$$c = \frac{V_{ub}^*}{A\lambda^3}, \quad \text{and} \quad b = \frac{V_{td}}{A\lambda^3}. \tag{9.103}$$

Both are complex numbers. $|V_{ub}|$ can be estimated from the decay $\bar{B}^0 \to \pi^+ \mu^- \bar{\nu}_\mu$, with a branching ratio $\sim 2 \times 10^{-4}$. As mentioned in Chapter 7, top quark elements like $|V_{td}|$ have been estimated from box diagrams that couple the top quark to the down quark, like in $B^0 \bar{B}^0$ mixing (see Figure 9.13). We have two different mass mixing matrices, the two component M of Equation (9.49) that describes weak mixing of neutral mesons, and the CKM matrix that describes

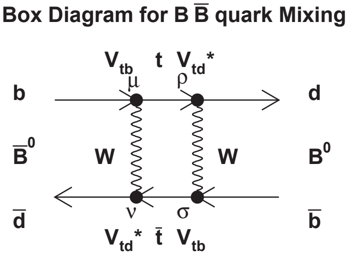

Box Diagram for B B̄ quark Mixing
V_{tb} t V_{td}*
μ
ρ
b
d
B̄⁰
W
W
B⁰
d̄
b̄
ν
σ
V_{td}* t̄ V_{tb}

Figure 9.13: Box diagram for $B\bar{B}$ mixing. The CKM matrix elements appear at the vertices to couple the quarks to the W's. Top quarks are the dominant contributors to $B\bar{B}$ mixing, but any charge $+2/3$ quark will work, and the top q and bottom $\bar{q}$ can be different flavors. The W's and the top quarks may be interchanged — same coupling. The Greek letters μ, ν, ρ, σ refer to indices of the γ matrices at each vertex. The first W propagator couples μ and ν. If the process is turned around, the CKM matrix elements are complex conjugated.

weak mixing of the charge $-1/3$ quarks. Neutral mesons can mix because there is a quantum number respected by strong and electromagnetic interactions that is violated by the weak interaction, meaning that particle and antiparticle can have common final states. There are no neutral quarks. Nature allows but suppresses weak transitions between quark doublets, and the CKM mixing matrix describes this phenomenon. M is Hermitian because it is a Hamiltonian operator, and CKM is unitary because it acts on wave functions, and has to preserve normalization.

The box diagrams shown in Figures 9.13 and 9.14 are quark model representations for the M_{12} mass mixing matrix element. For $B\bar{B}$ mixing, Figure 9.13 the top quark dominates the transition, so only top quarks are shown, although any charge 2/3 quark can be exchanged. The W's and the top quarks can be interchanged, so that quark lines flow on the two ends. That configuration shuffles the propagators, but involves the same product of CKM matrix elements at the vertices, namely $V_{tb}V_{td}^*V_{tb}V_{td}^*$. We adopt the convention of Ling Lie Chau [13] that for a $q_i\bar{q}_jW$ vertex, if the outgoing quark is u-like, the factor is V_{ij}, and if it is d-like it is V_{ij}^*. The matrix element acts on an outgoing charge 2/3 quark, and its complex conjugate acts on an outgoing charge 1/3 quark. $|V_{tb}| = 1$, and it is real, so the CP violating phase is contained in V_{td}^*. Feynman's rules

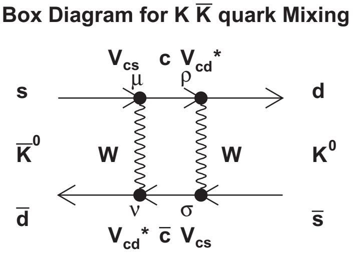

Box Diagram for K $\overline{K}$ quark Mixing
V_{cs} c V_{cd}^*
μ
ρ
s
d
$\overline{K}^0$
W
W
K^0
ν
σ
$\bar{d}$
$\bar{s}$
V_{cd}^* $\bar{c}$ V_{cs}

Figure 9.14: Box diagram for $K\bar{K}$ mixing. The CKM matrix elements appear at the vertices to couple the quarks to the W's. Charm quarks are the dominant contributors to $K\bar{K}$ mixing, but any charge +2/3 quark will work, and the top q and bottom $\bar{q}$ can be different flavors. The W's and the top quarks may be interchanged — same coupling. The Greek letters μ, ν, ρ, σ refer to indices of the γ matrices at each vertex. If the process is turned around, the CKM matrix elements are complex conjugated.

give the amplitude, but this diagram is more complicated than any we have seen so far, so it is worth going through it. There are four spinors — that is not unusual:

$$u_b(p_1), \quad \bar{v}_d(p_2), \quad \bar{u}_d(p_3), \quad \text{and} \quad v_b(p_4);$$
$$p_1 + p_2 = p_3 + p_4. \tag{9.104}$$

The index refers to the quark flavor (the spin index is omitted). There are four vertices; the letters μ, ν, ρ, σ refer to the indices of the gamma matrices at each vertex. Each vertex couples an on-shell quark and a quark propagator to a W propagator, and gets a factor:

$$-iV_{ij}\frac{g}{\sqrt{2}}\gamma^\mu\frac{(1-\gamma^5)}{2}, \quad \text{or} \quad -iV_{ij}^*\frac{g}{\sqrt{2}}\gamma^\rho\frac{(1-\gamma^5)}{2};$$
$$\text{top quark propagators are } i\frac{\not{k}+m_t}{(k^2-m_t^2)}. \tag{9.105}$$

Each W propagator carries the indices of the two vertices that it connects. For Figure 9.13, we have:

$$-i\frac{g_{\mu\nu} - k_\mu k_\nu/M_W^2}{k^2 - M_W^2}, \quad \text{and} \quad -i\frac{g_{\rho\sigma} - k_\rho k_\sigma/M_W^2}{k^2 - M_W^2}. \tag{9.106}$$

The four momentum k runs around the loop, and gets integrated over from $-\infty$ to $+\infty$. This is a new phenomenon for us — a loop integral! The integration has to be handled with care because of the singularities in the denominators of the propagators — something we have not had to worry about. We write down the amplitude for Figure 9.13:

$$\begin{aligned}
-M_{12} = \frac{1}{4}\mathcal{M}\sum_{\text{spins}}\left(-i\frac{g}{2\sqrt{2}}\right)^4 V_{tb}^2 V_{td}^{*2}\int\frac{d^4k}{(2\pi)^4} \\
\times\left\{\frac{g_{\mu\nu}-k_\mu k_\nu/M_W^2}{k^2-M_W^2}\frac{g_{\rho\sigma}-k_\rho k_\sigma/M_W^2}{k^2-M_W^2}\bar{u}_d(p_3)\gamma^\rho(1-\gamma^5)\right. \\
\times\frac{\not{k}+m_t}{k^2-m_t^2}\gamma^\mu(1-\gamma^5)u_b(p_1)\bar{v}_d(p_2)\gamma^\nu(1-\gamma^5) \\
\times\left.\frac{\not{k}+m_t}{k^2-m_t^2}\gamma^\sigma(1-\gamma^5)v_b(p_4)\right\}. \qquad (9.107)
\end{aligned}$$

The multiplier $\mathcal{M}$ in front is there to correct for the fact that the left-hand side is mixing for $B\bar{B}$ mesons, and the right-hand side is quarks. Evaluation of this integral is a chore, but it can be done, and the result is finite. Ling Lie Chau writes the result for $B\bar{B}$ in the following form [13]:

$$\begin{aligned}
M_{12} = \frac{G_F^2 m_W^2}{16\pi^2}m_B f_B^2 (V_{tb})^2 (V_{td}^*)^2\left\{\frac{x}{(1-x)^2}\left(3-\frac{19x}{4}+\frac{x^2}{4}\right)\right. \\
\left.+2\left(1-\frac{3x^2}{4(1-x)}\right)\frac{x\ln x}{(1-x)^2}\right\}, \qquad (9.108)
\end{aligned}$$

where $x = (m_t/M_W)^2 = 4.8$, and the function in brackets is 1.1. $\mathcal{M} = m_B f_B^2$, where f_B is the constant analogous to f_π and f_K discussed earlier, obtained from the transition rates for leptonic decays. Whether this product accurately corrects for quark binding or not can be checked by experiment. Equation (9.108) should equal $\Delta m_B/2$. A nice first check is the ratio:

$$\frac{\Delta m_{Bs}}{\Delta m_{Bd}} = \frac{V_{ts}^{*2}}{V_{td}^{*2}}\frac{m_{Bs}}{m_{Bd}}\frac{f_{Bs}^2}{f_{Bd}^2} = 21\frac{f_{Bs}^2}{f_{Bd}^2} = 35, \quad \frac{f_{Bs}}{f_{Bd}} = 1.3, \qquad (9.109)$$

in reasonable agreement with the PDG value $f_{Bs}/f_{Bd} = 1.21$. The magnitudes of the CKM matrix elements were obtained from Equation (9.90). Many factors cancel in this ratio. Equation (9.108) can be altered for $K\bar{K}$ mixing. In that case the charm quark replaces the top quark in the loop, giving $V_{cs}^2 V_{cd}^{*2}$ for the CKM matrix elements, and the curly bracket function is -3.4×10^{-3}, using $m_c = 1.3\,\mathrm{GeV}$. Then the ratio:

$$\frac{\Delta m_b}{\Delta m_K} = \frac{V_{td}^{*2}(1.1) m_B f_B^2}{V_{cd}^{*2}(3.4 \times 10^{-3}) m_K f_K^2} = \frac{5.5 f_B^2}{f_K^2} = 96, \qquad \frac{f_B}{f_K} = 4.2. \tag{9.110}$$

The PDG number for this ratio is 1.35, so the agreement is not so good. However, the K^0 is composed of two light quarks, $(d\bar{s})$, and the application of a quark diagram like Figure 9.14 is not reliable. On the other hand, if one were trying to reconcile the CKM matrix with the CP parameter ϵ in Equation (9.67), which is related to the imaginary part of M_{12}, then V_{cs} is of interest. This element was set to one in the above ratio, and is real in the Wolfenstein form to order λ^3, but not to order λ^4. If we knew better how to dress the light quarks, it would be possible to obtain ϵ from the CKM matrix, and explain all known CP violations in the quark sector with one complex phase. The phase to create ϵ was inserted by hand. The phase in the CKM matrix is also inserted by hand, but it includes light and heavy quarks, and predicts large CP violation effects in $B\bar{B}$ mixing.

In order to learn more about this, we return to the unitary triangle in the Wolfenstein approximation Figure 9.12. We focus on the angle β:

$$\sin\beta = \frac{\eta}{\sqrt{(1-\rho)^2 + \eta^2}}, \quad \cos\beta = \frac{1-\rho}{\sqrt{(1-\rho)^2 + \eta^2}},$$
$$\sin(2\beta) = 2\sin\beta\cos\beta = \frac{2\eta(1-\rho)}{(1-\rho)^2 + \eta^2}. \tag{9.111}$$

A relation among the CKM matrix elements and the angle β is

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad V_{tb} = 1, \quad V_{cd} = -\lambda,$$
$$V_{cb} = A\lambda^2, \quad V_{td} = A\lambda^3(1-\rho-i\eta). \tag{9.112}$$

Therefore

$$-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} = \frac{1-\rho+i\eta}{(1-\rho)^2+\eta^2} = \frac{e^{i\beta}}{\sqrt{(1-\rho)^2+\eta^2}}, \quad \text{where} \tan\,\beta = \frac{\eta}{1-\rho}. \tag{9.113}$$

So, the assertion in Equation (9.111) is correct. Another ratio of interest is

$$\frac{V_{td}^*}{|V_{td}|} = \frac{1-\rho+i\eta}{\sqrt{(1-\rho)^2+\eta^2}} = e^{i\beta}. \tag{9.114}$$

These formulas will prove useful in understanding the experiments that we are about to describe.

The angle β has been measured by studying CP violation in the $B\bar{B}$ system. The experiments were carried out at two different laboratories at approximately the same time, around the year 2000. They were based on the coherent source supplied by the $\Upsilon(4S) \to B\bar{B}$. The effect is a combination of mixing and CP violation, and is time dependent. The slowly separating $B\bar{B}$ pair from the $\Upsilon(4S)$ ($v = 0.061$, Equation (9.83)) made time dependent measurements impossible. So two laboratories, SLAC at Stanford, California, (PEPII) and KEK in Japan, (KEKB) constructed a second storage ring for the positrons, so that the collision could be asymmetric, with the center of mass moving in the laboratory. The ring energies at the two laboratories were slightly different, but approximately 10 GeV on 3 GeV. The kinematics is left as an exercise in Problem 9.12. The $\Upsilon(4S)$ (the center of mass) moves in the 10 GeV beam direction with $\gamma_c v_c = 0.425$ at KEKB, and 0.55 at PEPII. The $B^0\bar{B}^0$ pair moves together with this velocity. There is some transverse separation of the pair from the velocity in the center of mass, about 10% of v_c.

The mean decay length for PEPII is $\gamma v c\tau = 260\,\mu$m, and this is the average separation between two decay vertices. Remember that a senile muon lives just as long as one freshly made. The experiments at the two laboratories were named BaBar at PEPII and Belle at KEKB.

Each experiment has published numerous papers. The following discussion is based on BaBar [14] and Belle [11]. These two reports are very similar, with comparable integrated luminosity and sample size. The mixing formula, Equation (9.89), is a function of $|\Delta t|$, the time difference between the decays of the two B's. For completeness, we repeat it here, and add the formula for no oscillation. There is no mixing if there is no mass difference.

$$B\bar{B} \to BB/\bar{B}\bar{B}\ \frac{dPr}{dt} = \frac{\Gamma_B}{4} e^{-\Gamma_B|\Delta t|}(1 - \cos(\Delta m|\Delta t|)),$$

$$\int_0^\infty dPr = \frac{\Gamma_B}{4}\left(\frac{1}{\Gamma_B} - \frac{\Gamma_B}{\Gamma_B^2 + (\Delta m)^2}\right),$$

and

$$B\bar{B} \to B\bar{B}\ \frac{dPr}{dt} = \frac{\Gamma_B}{4} e^{-\Gamma_B|\Delta t|}(1 + \cos(\Delta m|\Delta t|)),$$

$$\int_0^\infty dPr = \frac{\Gamma_B}{4}\left(\frac{1}{\Gamma_B} + \frac{\Gamma_B}{\Gamma_B^2 + (\Delta m)^2}\right). \qquad (9.115)$$

The mass difference can be measured from the time integral. If from a known sample of $B\bar{B}$ pairs there are like sign charged lepton pairs, then the difference between the number of like sign vs opposite sign pairs divided by the sum is equal to $\Gamma^2/(\Gamma^2 + (\Delta m)^2) = 0.6$ for the B_d. For an upper time limit of $2/\Gamma_B$ about half of the B's have already become $\bar{B}'$s. The sum of the two integrals is one half, to account for the initial B decay being either B or $\bar{B}$. Measuring the time dependence of the oscillations is however convincing evidence of what is going on, and not possible for the $\Upsilon(4S)$ without an asymmetric B factory. Δt is converted to the B^0 rest frame from the difference in location of the two vertices along the common beam line.

The relation is

$$\Delta t = t - t_0 = \frac{(z_2 - z_1)}{c\gamma_c v_c} = \frac{\Delta z}{c\gamma_c v_c}, \quad \text{or } 100\,\mu\text{m} = 0.60\,\text{ps (BaBar)}. \tag{9.116}$$

The experimental vertex resolution depended on several factors, and improved with the number of charged tracks that could be associated with a particular vertex. The average Δz resolution for BaBar reconstructing $B\bar{B}$ pairs was 180 μm, corresponding to 1.08 ps or 0.69 τ_{B^0} in the B^0 rest frame. An additional experimental effect is the mistag rate, that is the probability that the decay was really a B^0, but was called a $\bar{B}^0$. If w is the mistag probability for a particular final state, then the cosine term is multiplied by a dilution factor $D = (1 - 2w)$, which decreases the sensitivity. Mistag rates can be obtained from the data by looking at events with small Δt, where the mixing probability is small, or alternatively assuming the world average value for Δm, simply obtain D directly from the oscillation amplitude. The vertex resolution of course affects the definition of $\Delta t = 0$. The time dependent probabilities Equation (9.115) are plotted in Figure 9.15. Unmixed and mixed are shown for $\Delta m = 0.5$/ps, with the mixing curve multiplied by 10 to show mixed and unmixed on the same plot. To demonstrate the sensitivity, the mixed curve for $\Delta m = 1.0$/ps is shown as well, not multiplied by 10. A larger mass difference gives an effect at earlier times, before suppression by the exponential decay. None of the curves show any clear oscillations. The lifetime has to be divided out for the cosine term to be obvious. The BaBar paper referenced above quotes a number for the mass difference from 6350 fully reconstructed B^0 decays to flavor eigenstates:

$$\Delta m_d = 0.516 \pm 0.016\ (\text{stat}) \pm 0.010\ (\text{sys})\,\text{ps}^{-1}. \tag{9.117}$$

A very nice result.

The mass difference comes from flavor tags: same sign vs opposite sign. The search for CP violation, η in Wolfenstein's form of the CKM matrix, was performed by looking at the rates for a flavor tag followed by a decay to a final state accessible to both B and $\bar{B}$, like $\pi^+\pi^-$ for K^0. The discussion around Equation (9.84) describes

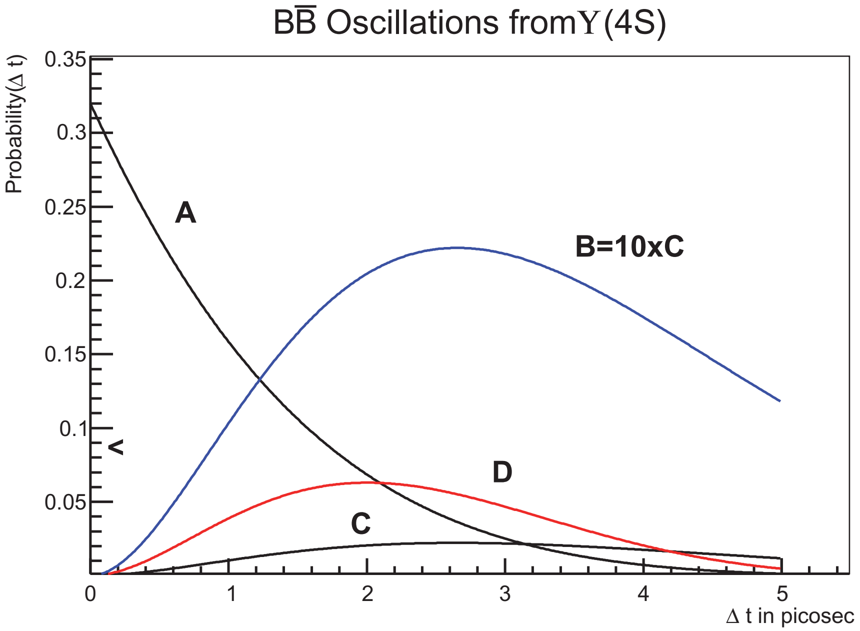


Figure 9.15: Plots of Equations (9.115) for no mixing curve A, mixing with $\Delta m = 0.5\,\mathrm{ps}^{-1}$ curve C, ×10 curve B (blue), and mixing with $\Delta m = 1\,\mathrm{ps}^{-1}$, curve D (red).

two of these final states, $J/\psi K_S$ CP odd, and $J/\psi K_L$ CP even. CP is violated if the rates for $B^0 \to J/\psi K_S$ and $\bar{B}^0 \to J/\psi K_S$ are not the same. In our treatment of CP violation in the K_S, K_L pair, the parameter ϵ comes from the mass mixing matrix, and the same argument holds for the B^0, so Equations (9.86) and (9.87), that described the mass mixing, have to be modified to include a CP violating parameter, call it ϵ_B. The normalization includes an extra factor. In analogy with Equations (9.67) and (9.69):

$$|B^0(t)\rangle = e^{-\Gamma t/2}\frac{\sqrt{2(1+|\epsilon_B|^2)}}{2(1+\epsilon_B)}(|B_L^0\rangle + |B_H^0\rangle\, e^{-i\Delta m t}),$$

$$|\bar{B}^0(t)\rangle = e^{-\Gamma t/2}\frac{\sqrt{2(1+|\epsilon_B|^2)}}{2(1-\epsilon_B)}(|B_L^0\rangle - |B_H^0\rangle\, e^{-i\Delta m t}), \quad (9.118)$$

with

$$\left|B_L^0\right\rangle = \frac{1}{\sqrt{2(1+|\epsilon_B|^2)}}((1+\epsilon_B)\left|B^0\right\rangle + (1-\epsilon_B)\left|\bar{B}^0\right\rangle),$$

$$\left|B_H^0\right\rangle = \frac{1}{\sqrt{2(1+|\epsilon_B|^2)}}((1+\epsilon_B)\left|B^0\right\rangle - (1-\epsilon_B)\left|\bar{B}^0\right\rangle). \tag{9.119}$$

The kets $\left|B_L^0\right\rangle$ and $\left|B_H^0\right\rangle$ (light and heavy) are normalized to one, but they are not orthogonal unless ϵ_B is pure imaginary: $\langle B_L|B_H\rangle = 2\mathrm{Re}(\epsilon_B)/(1+|\epsilon_B|^2)$. While we argued earlier that the box diagram for mass mixing did not give a reliable formula for ϵ in terms of CKM matrix elements, we now argue, based on the heavy mass of the b quark, that the box diagram *is* reliable for calculation of ϵ_B. Equation (9.75) reads:

$$\epsilon = \frac{Im(M_{12})}{2(\Gamma_{12}^2/4+|M_{12}|^2)}(-\Gamma_{12}/2+i|M_{12}|). \tag{9.120}$$

We have replaced M_{12} with $|M_{12}|$ in the denominator and the phase. It did not matter for K^0 because $\epsilon \sim 2\times 10^{-3}$, but ϵ_B is not small. The B_d^0 complex has a light and a heavy component, with the associated mass difference, but only one lifetime, meaning $\Gamma_{12} = 0$. Write $M_{12} = |M_{12}|e^{i\phi_{12}}$. Then:

$$\epsilon_B = \frac{i|M_{12}|^2\sin\phi_{12}}{2|M_{12}|^2} = i\sin\phi_{12}/2. \tag{9.121}$$

The phase angle ϕ_{12} comes from the CKM matrix elements in the box diagram for M_{12}. From Equation (9.107) for the box diagram and Equation (9.114), we have:

$$\phi_{12} = 2\beta, \tag{9.122}$$

where β is the angle of the unitary triangle between the real leg and $(1-\rho-i\eta)$. From Figure 9.12, it is clear that $\sin(2\beta)$ is of the order of unity, so ϵ_B is not small. It remains to work out formulas for $B^0 \to f$ and $\bar{B}^0 \to f$ as a function of time. Substituting (9.119) into (9.118) gives:

$$|B(t)\rangle = \frac{e^{-\Gamma t/2}}{2(1+\epsilon_B)}((1+\epsilon_B)(1+e^{-i\Delta m t})|B\rangle + (1-\epsilon_B)(1-e^{-i\Delta m t})|\bar{B}\rangle), \tag{9.123}$$

and

$$|\bar{B}(t)\rangle = \frac{e^{-\Gamma t/2}}{2(1-\epsilon_B)}((1+\epsilon_B)(1-e^{-i\Delta m t})|B\rangle + (1-\epsilon_B)(1+e^{-i\Delta m t})|\bar{B}\rangle). \tag{9.124}$$

Then

$$\langle f|B(t)\rangle = \frac{e^{-\Gamma t/2}}{1+\epsilon_B}\langle f|B\rangle(1+\epsilon_B e^{-i\Delta m t}), \tag{9.125}$$

and

$$\langle f|\bar{B}(t)\rangle = \frac{e^{-\Gamma t/2}}{1-\epsilon_B}\langle f|B\rangle(1-\epsilon_B e^{-i\Delta m t}). \tag{9.126}$$

We have assumed that $\langle f|B\rangle = \langle f|\bar{B}\rangle$. The validity of this assumption comes from the quark diagrams in Figure 9.16, if the final state $|f\rangle$ is CP even. If $|f\rangle$ is CP odd, the sign of the effect flips.

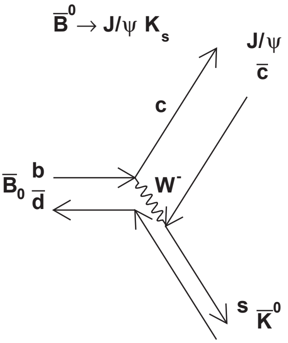


Figure 9.16: Quark flow diagram for $\bar{B}^0 \to J/\psi K_S$. The $\bar{d}$ quark is a spectator. The weak decay, with the CKM matrix element V_{cb} is $b \to c + W^-$. Then $W^- \to \bar{c} + s$ with V_{cs}^*. The K_S contains both K^0 and $\bar{K}^0$, so the charge conjugate diagram also works.

For $B^0 \to J/\psi K_S$, the product is $V_{cb}^* V_{cs}$, while for $\bar{B}^0 \to J/\psi K_S$ it is the complex conjugate $V_{cb} V_{cs}^*$. In the Wolfenstein approximation, Equation (9.91), both V_{cb} and V_{cs} are real (the η term in these matrix elements is small, $\sim\lambda^4$), which verifies the assumption to order λ^3. For CP odd, $\langle f|B\rangle = -\left\langle f\middle|\bar{B}\right\rangle$, and the result in Equation (9.127) changes sign. Verification is left as an exercise in Problem (9.15). Now $\epsilon_B = i\sin(2\beta)/2$ is pure imaginary, so the CP asymmetry, the difference over the sum, is

$$\begin{aligned}\frac{|\langle f|B(t)\rangle|^2 - |\left\langle f\middle|\bar{B}(t)\right\rangle|^2}{|\langle f|B(t)\rangle|^2 + |\left\langle f\middle|\bar{B}(t)\right\rangle|^2} &= \frac{2\mathrm{Re}(\epsilon_B e^{-i\Delta mt})}{1+|\epsilon_B|^2} \\ &= \frac{\sin(2\beta)\sin(\Delta mt)}{1+\sin^2(2\beta)/4}. \qquad (9.127)\end{aligned}$$

This is for CP even, like $J/\psi K_L$. For CP odd, $J/\psi K_S$, the sign of the exponent changes and that flips the answer to $-\sin(2\beta)\sin(\Delta mt)$.

Figure 9.17 shows the results of the Belle Collaboration on the measurement of the time dependent asymmetry Equation (9.126) for CP even and CP odd events. From the amplitude of the oscillation they report a result:

$$\begin{aligned}\sin(2\beta) &= 0.99 \pm 0.14\,(\mathrm{stat}) \pm 0.06\ (\mathrm{sys}); \\ \text{2020 PDG world average}\ \sin(2\beta) &= 0.698 \pm 0.027. \qquad (9.128)\end{aligned}$$

The Belle collaboration concluded in *Phys. Rev. Lett.* that CP violation had been observed in the $B^0\bar{B}^0$ system. There were 3×10^7 $\Upsilon(4S)$, so in that sense the effect was small. But for the common final CP eigenstates the effect is nearly 100%. Similar results were obtained simultaneously by BaBar, and since then experiments have continued, using both the $\Upsilon(4S)$ and hadronic $B\bar{B}$ production. Work is ongoing at the LHC at CERN and at KEK in Japan. There is an ambiguity in obtaining β from $\sin(2\beta) = 2\sin\beta\cos\beta$. From $\sin(2\beta) = 0.7$, $2\beta = 44^\circ$, or $2\beta = 136^\circ$, and $\beta = 22^\circ$ or $\beta = 68^\circ$. The smaller value fits better on the diagram.

We end the text on a positive note. The observation of large CP violating effects in the decay of heavy quarks is a remarkable

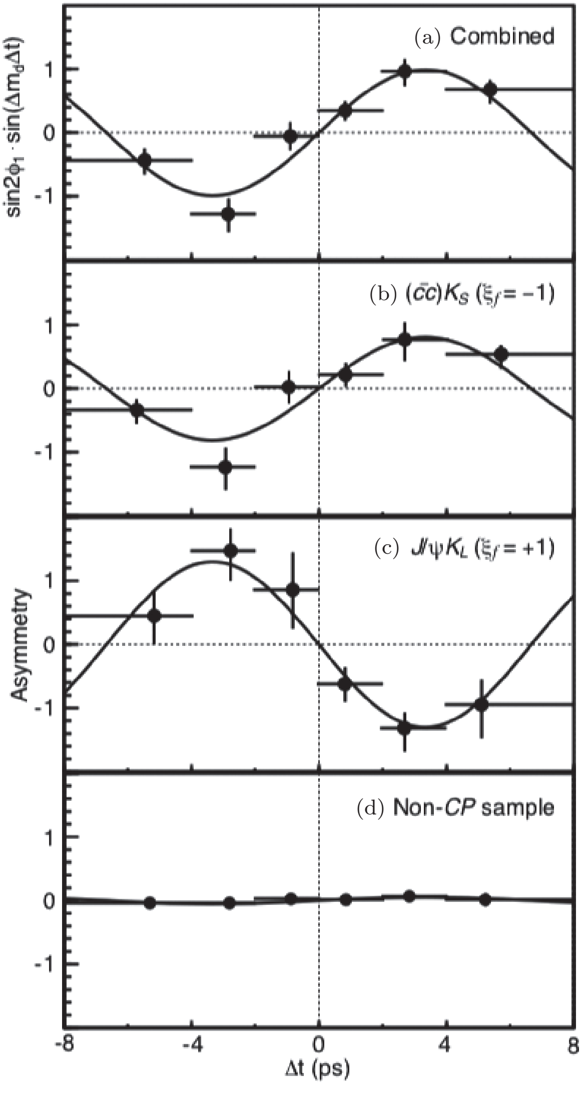


Figure 9.17: Results for measurement of $\sin(2\beta)$, called $\sin(2\phi_1)$ in the figure, from a sample of 747 CP odd and 569 CP even events from $\Upsilon(4S) \rightarrow B^0\bar{B}^0$. The sign of the effect has been changed to combine the two samples in the top curve, while the odd and even samples are shown separately in the next two curves. There is a large background in the CP even sample (the K_L), but the background does not oscillate! Credit Ref. [10].

confirmation of the quark model and the CKM mixing matrix. Further advances in our knowledge should explain why the CKM matrix is what it is, and where η comes from. For now, we must be satisfied with the success of the phenomenology. Other subjects of current interest include heavy exotic mesons that do not fit $q\bar{q}$, top quarks, and Higgs. The CKM matrix allows CP violation for the quarks, but makes no statement about the leptons. Neutrino physics has hardly been scratched in this book. Indeed, there are whole books devoted to this subject, and there are active research programs in neutrino mixing and CP violation today. Elementary particle physics is a mature, sophisticated subject, and many natural phenomena can be mathematically described by the standard model. The reader of this book has been exposed to techniques of calculation that are accurate and yet accessible to many people with some familiarity of quantum mechanics.

9.8. Problems

9.1 The classical form of the Lenz vector defined for $L \neq 0$ is

$$\vec{M} = \frac{1}{m}\vec{p} \times \vec{L} - \frac{e^2\vec{r}}{r}. \tag{9.129}$$

This is a constant vector in the plane of the orbit that is parallel to the semi-major axis of the ellipse, the line containing the foci. It represents the fact that the orientation of the ellipse remains fixed in space. For a $1/r^2$ central force there is no ‘precession of the perihelion’. Show that $d\vec{M}/dt = 0$, if $d\vec{L}/dt = 0$ and $d\vec{p}/dt = -e^2\vec{r}/r^3$. The classical observables commute! The quantum mechanical observable is symmetrized to be Hermitian, because $(\vec{L} \times \vec{p})^\dagger = -\vec{p} \times \vec{L}$:

$$\begin{aligned}\vec{M} &= \frac{1}{2m}(\vec{p} \times \vec{L} - \vec{L} \times \vec{p}) - \frac{e^2\vec{r}}{r} \\ &= \frac{1}{m}(\vec{p} \times \vec{L}) - \frac{i\hbar\vec{p}}{m} - \frac{e^2\vec{r}}{r}.\end{aligned} \tag{9.130}$$

$M^\dagger = M$. Some useful formulas are

$$\vec{L}\cdot\vec{M} = 0, \quad \vec{L}\times\vec{p}+\vec{p}\times\vec{L} = 2i\hbar\vec{p},$$

$$\vec{r}\cdot\vec{p}\times\vec{L} = \vec{L}^2, \text{ interchange dot and cross,}$$

and

$$[L_k, p_l] = i\hbar\epsilon_{klm}p_m, \quad \vec{p}\cdot\vec{L} = 0, \tag{9.131}$$

so that

$$(\vec{p}\times\vec{L})^2 = \sum \epsilon_{ijk}\epsilon_{ilm}p_j L_k p_l L_m = \vec{p}^{\,2}\vec{L}^2. \tag{9.132}$$

$\vec{M}$ is a constant of the motion, $[\vec{M}, H] = 0$, so the Lenz vector in quantum mechanics represents a degeneracy. Evaluation of the commutator is complicated because $\vec{p}$ and $\vec{r}$ do not commute. Use the definition of time dependence of an operator:

$$\frac{dM_x}{dt} = \frac{1}{i\hbar}[M_x, H] = 0 = \frac{1}{m}\left\{\frac{d(\vec{p}\times\vec{L})_x}{dt} - \frac{i\hbar dp_x}{dt}\right\} - e^2\frac{d}{dt}\frac{x}{r},$$

$$H = \frac{\vec{p}^{\,2}}{2m} - \frac{e^2}{r}. \tag{9.133}$$

Show that $1/m$ times the term in brackets is:

$$\frac{e^2}{mr^3}(r^2p_x - x(\vec{r}\cdot\vec{p} - i\hbar)). \tag{9.134}$$

Then show that $e^2 d/dt(x/r)$ is the same, proving that the commutator vanishes.

9.2 Work out the quantum numbers in spherical coordinates for the third excited state of the three dimensional isotropic harmonic oscillator: $n_x + n_y + n_z = 3$. Match degeneracy, parity, and E_n. The radial wave function for a three-dimensional harmonic oscillator may be written, following Problem 10.10 in Merzbacher:

$$\psi(r,\theta,\phi) = r^l e^{-\alpha^2 r^2/2} f(r) Y_l^{m_l}(\theta,\phi),$$

$$\alpha^4 = \frac{mk}{\hbar^2}, \quad \omega^2 = \frac{k}{m}. \tag{9.135}$$

The radial wave equation is

$$\left\{\frac{d^2}{dr^2} + \frac{2d}{rdr} - \left(\frac{mkr^2}{\hbar^2} + \frac{l(l+1)}{r^2} - \frac{2mE}{\hbar^2}\right)\right\} R(r) = 0. \tag{9.136}$$

Substitute $R(r)$ from Equation (9.135) to obtain:

$$f''(r) + \left(\frac{2(l+1)}{r} - 2\alpha^2 r\right) f'(r) + 4\alpha^2 \left(\frac{E}{2\hbar\omega} - \frac{(2l+3)}{4}\right) f(r) = 0. \qquad (9.137)$$

The Gaussian asymptotic form takes out the harmonic oscillator potential, and r^l takes out the centrifugal potential. Hydrogen atom wave functions resemble Equation (9.135), except that a damped exponential replaces the Gaussian. This is a clue that a change of variables $r \to r^2$ is needed to relate the two problems. This is what is meant in the text by 'judicious choice of variables'. Associated Laguerre polynomials $L^p_{q-p}(z)$ satisfy the differential equation:

$$z\frac{d^2w}{dz^2} + (p+1-z)\frac{dw}{dz} + (q-p)w = 0. \qquad (9.138)$$

Replace $z \to \alpha^2 r^2$ and obtain the form:

$$\frac{d^2w}{dr^2} + \left(\frac{2p+1}{r} - 2\alpha^2 r\right)\frac{dw}{dr} + 4\alpha^2(q-p)w = 0. \qquad (9.139)$$

Comparison of Equations (9.137) and (9.139) shows that $p = l + 1/2$ and $q - p = n - 1$, where $E_n = \hbar\omega(2n + l - 1/2)$. The indices q and p are half integers, but $q - p$ is an integer. The solutions are generalizations of associated Laguerre polynomials for half integer p. This is the connection between hydrogen and the harmonic oscillator.

9.3 Contribution of the annihilation diagram $e^-e^+ \to \gamma \to e^-e^+$ to the positronium ground state. First we need a relation between the relativistic T matrix element calculated by the Feynman rules, and a non-relativistic potential $V(r)$. The link is made via the Born approximation for scattering. Standard quantum mechanics texts (Merzbacher, Chapter 11, Section 4) give the scattering cross-section in terms of the potential and the momentum transfer in the center of mass as

$$\frac{d\sigma}{d\Omega} = \frac{(m_r c^2)^2}{4\pi^2(\hbar c)^4}\left|\int e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}V(\vec{r})d^3r\right|^2. \tag{9.140}$$

For Bhabha scattering the reduced mass $m_r = m_e/2$. Compare this to Equation (4.69) in cgs units:

$$\frac{d\sigma}{d\Omega} = \frac{\hbar^2 c^2}{64\pi^2 s}|\langle f|\,T\,|i\rangle|^2. \tag{9.141}$$

In the non-relativistic limit for Bhabha scattering $s = (2m_e c^2)^2$. Comparing these two equations gives:

$$\int e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}V(\vec{r})d^3r = \frac{(\hbar c)^3}{4(m_e c^2)^2}\langle f|\,T\,|i\rangle. \tag{9.142}$$

The potential is given by the Fourier transform:

$$V(r) = \frac{(\hbar c)^3}{4(m_e c^2)^2}\frac{1}{(2\pi)^3}\int d^3q e^{-i\vec{q}\cdot\vec{r}}\langle f|\,T\,|i\rangle, \quad \vec{q} = \vec{k} - \vec{k}'. \tag{9.143}$$

Here, it is understood that the T matrix element is evaluated in the non-relativistic limit. The sign of the T matrix element now becomes important, because the potential can be either attractive or repulsive. The Feynman rules keep track of whether the force is attractive $T < 0$, or repulsive $T > 0$. This can be demonstrated in the non-relativistic approximation by relating T to H (see Equation (4.39) in Sakurai '*Advanced Quantum Mechanics*'). Cross-sections in first-order covariant perturbation theory, like the Born approximation, are independent of the sign of the charge, but this is not true for higher-order corrections, and the cross-sections for e^- and e^+ on high Z nuclei are quite different.

Evaluate T in the non-relativistic limit for Bhabha scattering by single photon exchange:

$$\langle f|\,T\,|i\rangle = -\frac{e^2}{q^2}\bar{u}(p_3)\gamma^\mu u(p_1)\bar{v}(p_2)\gamma_\mu v(p_4). \tag{9.144}$$

Obtain the non-relativistic limit averaged over initial spins and summed over final spins:

$$\frac{1}{4}\sum_{\text{spins}} \langle f|\, T\, |i\rangle = -\frac{4e^2m_e^2}{q^2} = -\frac{16\pi(m_ec^2)^2\alpha}{\hbar^2c^2q^2}, \tag{9.145}$$

where we have used the fine structure constant to help convert from natural units to cgs units, and $\vec{q}$ is now a wave number, with dimensions $1/l$. The T matrix element is still dimensionless, and $\alpha = e^2/(\hbar c) = 1/137$. Substitute this in Equation (9.143), and use $\int_0^\infty dx(\sin x)/x = \pi/2$ to obtain:

$$V(r) = -\frac{e^2}{r}. \tag{9.146}$$

So this works. Now apply the same analysis to the annihilation diagram:

$$\langle f|\, T\, |i\rangle = -\frac{e^2}{q^2}\bar{v}(p_2)\gamma^\mu u(p_1)\bar{u}(p_3)\gamma_\mu v(p_4). \tag{9.147}$$

There is a sign difference that comes from the matching of e^- and e^+ spinors, and replacing γ^0 with $\vec{\gamma}$. The other important difference is that the propagator is now a constant: $q^2 = 4m_e^2$, and the effective potential is a $\delta(\vec{r})$ function. Show that only triplet states contribute, and obtain the result for the expectation value of the annihilation potential for the positronium ground state wave function:

$$\langle V_{ann}\rangle = \frac{(\hbar c)^3\pi\alpha}{(m_ec^2)^2}|\psi_{n=1,l=0}(0)|^2 = \frac{m_ec^2\alpha^4}{4}. \tag{9.148}$$

9.4 This problem builds on the previous one to calculate the singlet–triplet splitting of the positronium ground state. The annihilation term Equation (9.148) gets added to the usual magnetic moment interaction term. The vector potential from a classical magnetic moment $\vec{m}$ at the origin is given in cgs units by the formula:

$$\vec{A} = -\vec{m}\times\vec{\nabla}\left(\frac{1}{r}\right), \quad \text{so } \vec{B} = -\vec{\nabla}\times\vec{m}\times\vec{\nabla}\left(\frac{1}{r}\right). \tag{9.149}$$

The classical magnetic moment is replaced by the positron magnetic moment operator $\vec{\mu}_{pos} = e\hbar/(2m_e c)\vec{\sigma}_2$ in the quantum mechanical Hamiltonian. For the positronium hyperfine structure we write:

$$H_{\text{int}} = -\vec{\mu}_{\text{ele}} \cdot \vec{B} = -\left(\frac{e\hbar}{2m_e c}\right)^2 \vec{\sigma}_1 \cdot (\vec{\nabla} \times \vec{\sigma}_2 \times \vec{\nabla}) \left(\frac{1}{r}\right). \tag{9.150}$$

The two magnetic moments are equal in magnitude and opposite in sign. The triple cross product can be rewritten using the usual formula $\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$, which is the same as $\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}$. Thus,

$$H_{\text{int}} = -\left(\frac{e\hbar}{2m_e c}\right)^2 (\vec{\sigma}_1 \cdot \vec{\sigma}_2 \vec{\nabla}^2 - (\sigma_1 \cdot \vec{\nabla})(\sigma_2 \cdot \vec{\nabla}))\frac{1}{r}, \tag{9.151}$$

and the second term is

$$-(\vec{\sigma}_1 \cdot \vec{\nabla})(\vec{\sigma}_2 \cdot \vec{\nabla}) \left(\frac{1}{r}\right) = -\left(\sigma_{1x}\frac{\partial}{\partial x} + \sigma_{1y}\frac{\partial}{\partial y} + \sigma_{1z}\frac{\partial}{\partial z}\right) \times \left(\sigma_{2x}\frac{\partial}{\partial x} + \sigma_{2y}\frac{\partial}{\partial y} + \sigma_{2z}\frac{\partial}{\partial z}\right)\left(\frac{1}{r}\right). \tag{9.152}$$

Equation (9.152) has to be evaluated carefully. ‘Diagonal’ terms like $(\sigma_{1x}\sigma_{2x}\partial^2/\partial x^2)(1/r) = (1/3)\sigma_{1x}\sigma_{2x}\vec{\nabla}^2(1/r)$, so the sum of diagonal terms subtracts 1/3 from the first term. The space parts of off diagonal terms like $\sigma_{1x}\sigma_{2y}\partial^2/\partial x\partial y(1/r)$ transform like the Y_2^m spherical harmonics. Since we are interested in the ground state of positronium that has $l = 0$, we subtract $(\vec{\sigma}_1 \cdot \vec{\sigma}_2\vec{\nabla}^2)/3$ from the first term, and then ignore what remains of the second part.

$$H_{\text{int}} = -\frac{2}{3}\left(\frac{e\hbar}{2m_e c}\right)^2 (\vec{\sigma}_1 \cdot \vec{\sigma}_2)\vec{\nabla}^2 \left(\frac{1}{r}\right). \tag{9.153}$$

$\vec{\nabla}^2(1/r) = -4\pi\delta(\vec{r})$ is the singularity referred to in the text. Evaluate the singlet–triplet splitting using the Hamiltonian in

Equation (9.153), and combine it with the annihilation diagram Equation (9.148) to obtain the final result:

$$\begin{aligned}\Delta E_{hfs} &= E(n=1, l=0, \text{triplet}) - E(n=1, l=0, \text{singlet}) \\ &= \left(\frac{1}{3} + \frac{1}{4}\right) \alpha^4 m_e c^2. \qquad (9.154)\end{aligned}$$

Give the number in eV and in MHz. This number can be accurately measured. Early experiments have been described by Deutsch [16]. About 30% of positrons stopping in a gas form positronium. The rest annihilate in flight. The triplet/singlet ratio is 3/1. Triplet positronium decays into three γ rays that give a continuous energy spectrum, while the singlet state gives a line spectrum at 511 keV. Thus, microwaves of the correct frequency will cause the triplet to singlet transition, and a decrease in the continuous γ spectrum. Deutsch's original measurements were made in a magnetic field where the Zeeman splitting between $J = 1, m = \pm 1$ and $J = 1, m = 0$ was measured — a smaller energy difference than Equation (9.154). The reason for application of the Zeeman effect was that at that time microwave sources in the 1 mm wavelength region were not available, but in the 2 cm range they were.

9.5 Zeeman effect in positronium. Positronium has the curious feature that it has no linear Zeeman effect. That means that it has no permanent magnetic dipole moment. Since e^- and e^+ chase each other around the orbit, there is no current associated with the angular momentum $\vec{L}$, and no orbital magnetic moment. There is no spin magnetic moment either. Obviously for $S = 1$ there is no moment, because if $m_1 = +1/2$ and $m_2 = +1/2$ the magnetic moments cancel, and there is no net effect for $m_s = 0$. The effect is for $m_1 = +1/2$ and $m_2 = -1/2$ or vice versa, decoupling the total S eigenstates. The ground state interaction term in a magnetic field is

$$H_B = -(\vec{\mu}_+ + \vec{\mu}_-) \cdot \vec{B} = -\frac{|e|}{mc}(\vec{S}_+ - \vec{S}_-) \cdot \vec{B}. \qquad (9.155)$$

Let m_1 be the electron and m_2 the positron. Then it is straightforward to calculate the matrix elements:

$$\langle m_1 = +1/2,\ m_2 = -1/2|\, H_B \,|m_1 = +1/2,\ m_2 = -1/2\rangle$$
$$= \frac{|e|\hbar}{mc} B_z = \Delta_B,$$

and

$$\langle m_1 = -1/2,\ m_2 = +1/2|\, H_B \,|m_1 = -1/2,\ m_2 = +1/2\rangle$$
$$= -\Delta_B. \tag{9.156}$$

Diagonalize the matrix in m_1, m_2 space combining Equations (9.155) and (9.156). It is a 2×2 matrix, because $+1/2$, $+1/2$ and $-1/2$, $-1/2$ are decoupled. Obtain the eigenvalues:

$$\lambda_\pm = (E_1 + E_0)/2 \pm \sqrt{(E_1 - E_0)^2/4 + \Delta_B^2}, \tag{9.157}$$

where $E_1 - E_0 = \Delta E_{hfs}$, and E_1 is the $S = 1$ energy level.

Expand the square root for small $\Delta_B/\Delta E_{hfs}$, and show that the $S = 1, m_s = 0$ level and the $S = 0, m_s = 0$ level spread apart as a quadratic function of the magnetic field. The original measurement of ΔE_{hfs} was done using the formula for λ_+, and exploiting the fact that $S = 1$ $m_s = 0$ decoupled into ↑↓ and ↓↑ states causes $S = 0$ to appear, with the resulting two γ decay. The rf frequency was fixed, and the magnetic field varied to observe a decrease in the 3γ decays, that have a continuous energy spectrum (2γ decays are all at 511 keV). The frequency was then inversely proportional to ΔE_{hfs}, and the numerator Δ_B was known.

9.6 Muonium precession in an external magnetic field. Assume the muon is 100% polarized along its direction of motion. When the muon stops in material and captures an electron, half of the electron spins will be parallel to the muon spin, and half antiparallel. The antiparallel spins will form atoms with $m_s = 0$, so half of the muons will be depolarized. The other half will form muonium with $S = 1$, and $\langle S_y \rangle = +\hbar$ at $t = 0$. Show that the spin motion Hamiltonian is:

$$H = -\vec{\mu} \cdot \vec{B} = -\frac{eB}{m_e c}\left(-S_z + \left(1 + \frac{m_e}{m_\mu}\right) S_{2z}\right), \tag{9.158}$$

where $\vec{S} = \vec{S}_1 + \vec{S}_2$, and S_2 is the muon. Using the commutation relations, show that

$$\begin{aligned} \frac{dS_y}{dt} &= \frac{eB}{2m_e c}\left(1 - \frac{m_e}{m_\mu}\right) S_x, \quad \text{and} \\ \frac{dS_x}{dt} &= -\frac{eB}{2m_e c}\left(1 - \frac{m_e}{m_\mu}\right) S_y. \end{aligned} \tag{9.159}$$

Hence the precession frequency is $\omega = (eB)/(2m_e c)(1 - m_e/m_\mu)$, about 100 times the muon frequency. The magnetic moment is essentially that of the electron, but the angular momentum vector is twice as large.

9.7 Rate for muonium annihilation: $\mu^+ e^- \to \bar{\nu}_\mu \nu_e$. This is similar to $e^- e^+ \to \gamma\gamma$, only using the weak rather than the electromagnetic interaction. There is only one W t channel exchange diagram, because the Z does not couple the muon to an electron, so the problem is simpler than positronium. The momenta are: electron $p_1 = (m_e, 0, 0, 0)$; muon $p_2 = (m_\mu, 0, 0, 0)$; ν_e $p_3 = (m_\mu/2, 0, 0, m_\mu/2)$; and $\bar{\nu}_\mu$ $p_4 = (m_\mu/2, 0, 0, -m_\mu/2)$. The total energy is really $\sqrt{s} = m_e + m_\mu$, but we dropped the electron mass, and we chose the z axis for the two final state neutrino momenta. Check that the Feynman rules give the T matrix element:

$$\begin{aligned} -i \langle \nu_e \cdot \bar{\nu}_\mu | T \left| e^- \mu^+ \right\rangle & \\ = \bar{u}(p_3) \frac{-ig}{\sqrt{2}} \gamma^\mu \frac{1 - \gamma^5}{2} u(p_1) & \frac{-i(g_{\mu\nu} + q_\mu q_\nu / M_W^2)}{q^2 - M_W^2} \\ \times \bar{v}(p_2) \frac{-ig}{\sqrt{2}} \gamma^\nu \frac{1 - \gamma^5}{2} v(p_4). & \end{aligned} \tag{9.160}$$

$|q^2| = |(p_1 - p_3)^2| \ll M_W^2$, so replace the W propagator with $ig_{\mu\nu}/M_W^2$, and then use $G_F/\sqrt{2} = g^2/(8M_W^2)$ to introduce the Fermi coupling constant into the matrix element. Carefully write down the four spinors, using the facts that the electron and muon are at rest, and the muon is an antiparticle spinor. The neutrino–antineutrino are not at rest, but they are massless Dirac particle–antiparticle. Show that only γ^0 contributes to the matrix element, and use $\bar{u}\gamma^0 = u^\dagger$ to simplify the equation. Write down the 4×4 matrix $(1 - \gamma^5)$, and show that at each vertex the lepton and its neutrino have parallel spins, and that the ν_e is left handed, and the $\bar{\nu}_\mu$ right handed, so the muonium is in a triplet spin state. You knew that this had to be, but it is fun to see it actually work. Substitute the matrix element into the proper form of Equation (9.28), calculate $\langle B|B\rangle$. The muonium wave function is the same as hydrogen for our purposes. Evaluate the ratio of transition rates:

$$\frac{\Gamma(\mu^+e^- \to \bar{\nu}_\mu \nu_e)}{\Gamma(\mu^+ \to e^+\bar{\nu}_\mu \nu_e)} = \left(\frac{\alpha m_e}{m_\mu}\right)^3 \times 12\pi. \qquad (9.161)$$

This is a very small number, but an instructive exercise.

9.8 Do a literature search for experimental evidence for mesons and baryons containing more than one heavy quark. Start with the Particle Data Group and references therein. This is an active field of research, with new results from the LHC detectors every year.

9.9 Verify Equation (9.39), and try to fit it to $1^3S - 2^3S - 3^3S$ splittings for Υ and ψ families. Work in MeV-fermi units. $\hbar c = 197\,\text{MeV-fm}$, and let $r_0 = 0.5\,\text{fm}$, for the radius where the potential crosses the axis, although you do not really need either number to calculate energy level differences. The two families would be identical if Equation (9.39) were correct, and if $\Upsilon(3S) - \Upsilon(1S) = 895\,\text{MeV}$, then the prediction for $\Upsilon(3S) - \Upsilon(2S)$ is about $20\,\text{MeV}$ low, not too bad. You have only one free parameter, the constant C in the potential $V(r) = C\ln(r/r_0)$.

9.10 Verify the various relations listed in Equation (9.68). In particular obtain the formulas for p^2 and q^2.

9.11 Verify Equation (9.45) for $\vec{\sigma} \cdot \vec{r}/r$ in terms of spherical harmonics, and use this form to show that $\langle s = 0|\, S_{12}\, |s = 0\rangle =$

0, $\langle s=1, m_s=\pm 1|\, S_{12}\, |s=1, m_s=\pm 1\rangle \;=\; 3\cos^2\theta - 1$, and $\langle s=1, m_s=0|\, S_{12}\, |s=1, m_s=0\rangle = 2(1-3\cos^2\theta)$.

9.12 Plot Equation (9.82) for the interference between regenerated $K_S \to \pi^+\pi^-$ and CP violating $K_L \to \pi^+\pi^-$. Estimate the relative normalization of the interference term from the experimental plot Figure 9.11. On the same graph plot the curve for an anti-carbon regenerator, demonstrating clearly how matter versus antimatter can be determined.

9.13 The Wolfenstein form for the CKM matrix to order λ^5 for imaginary terms is [12]:

$$V_{\text{CKM}}$$
$$= \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta + i\eta\lambda^2/2) \\ -\lambda & 1-\lambda^2/2 - i\eta A^2\lambda^4 & A\lambda^2(1+i\eta\lambda^2) \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{9.162}$$

Note that now V_{cs} and V_{cb} have acquired complex phases that are order λ^4, and a very small part $\sim\lambda^5$ has been added to V_{ub}. The three independent unitary triangles are Equations (9.99)–(9.101).

$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0,$$
$$V_{ud}^* V_{us} + V_{cd}^* V_{cs} + V_{td}^* V_{ts} = 0,$$
$$V_{us}^* V_{ub} + V_{cs}^* V_{cb} + V_{ts}^* V_{tb} = 0.$$

Top is the golden triangle. The other two are flat in Wolfenstein's $\sim\lambda^3$ approximation. Use the more accurate form Equation (9.162) to plot all three triangles in a two dimensional complex space. Show that the second and third triangles, like the first, are almost exactly right angle triangles, and that all three have the same area.

9.14 Work out the kinematics for an asymmetric e^-e^+ collider tuned to the $\Upsilon(4S)$. Set the electron ring at 10 GeV. Calculate the energy of the positron ring, and the velocity of the $\Upsilon(4S)$ in the laboratory. What is the maximum opening angle and spatial distance between the two vertices B and $\bar{B}$ if they decay simultaneously after traveling one lifetime?

9.15 Given that $\langle B|\bar{B}\rangle = 0$, and $\langle B|B\rangle = \langle \bar{B}|\bar{B}\rangle = 1$, show that Equations (9.123) and (9.124) are correctly normalized at $t = 0$.

9.16 Verify the relation for the $B\bar{B}$ mass mixing matrix element $\mathrm{Im}(M_{12}) = |M_{12}|\sin(2\beta)$.

9.17 Show that the numerator and denominator for Equation (9.127) are

$$|\langle f|B(t)\rangle|^2 - |\langle f|\bar{B}(t)\rangle|^2 = \frac{|\langle f|B\rangle|^2 e^{-\Gamma t}}{|1+\epsilon_B|^2} 4\mathrm{Re}(\epsilon_B e^{-i\Delta m t}), \tag{9.163}$$

and

$$|\langle f|B(t)\rangle|^2 + |\langle f|\bar{B}(t)\rangle|^2 = \frac{|\langle f|B\rangle|^2 e^{-\Gamma t}}{|1+\epsilon_B|^2} 2(1+|\epsilon_B|^2). \tag{9.164}$$

9.18 Show that if $\langle f|B\rangle = -\langle f|\bar{B}\rangle$ a minus sign appears in Equation (9.127).

References

[1] L. D. Landau, *Doklady Akademii Nauk SSSR* **60**, 207 (1948).

[2] E. D. Theriot Jr. *et al.*, *Phys. Rev. A* **2**, 707 (1970).

[3] V. W. Hughes *et al.*, *Phys. Rev. Lett.* **5**, 63 (1960).

[4] S. H. Aronson *et al.*, *Phys. Rev. Lett.* **48**, 1078 (1982).

[5] S. W. Herb *et al.*, *Phys. Rev. Lett.* **39**, 252 (1977).

[6] C. Quigg and J. L. Rosner, *Phys. Reports* **56**, 167 (1979).

[7] J. Schwinger, *Quantum Mechanics* (Springer, Berlin, 2003), Equation 8.3.28.

[8] W. Carithers, *et al.*, *Phys. Rev. Lett.* **34**, 1240 and 1244 (1975).

[9] T. D. Lee, R. Oehme, and C. N. Yang, *Phys. Rev.* **106**, 340 (1957).

[10] J. Christenson, J. W. Cronin, V. Fitch, and R. Turlay, *Phys. Rev. Lett.* **13**, 138 (1964).

[11] K. Abe *et al.* (Belle Collaboration), *Phy. Rev. Lett.* **87**, 091802 (2001).

[12] L. Wolfenstein, *Phy. Rev. Lett.* **51**, 1945 (1983).

[13] L. L. Chau, *Phys. Reports* **95**, 1 (1983).

[14] B. Aubert *et al.*, *Phy. Rev. D* **66** 032003 (2002).
[15] D. Besson and T. Skwarnicki, *Ann. Rev. Part. Sci.* **43**, 333 (1993).
[16] M. Artuso *et al.* (Cleo Collaboration), *Phys. Rev. Lett.* **94**, 032001 (2005).
[17] M. Deutsch, *Proc. Ame. Acad. Arts Sci.* **82**, 331 (1953).

This page intentionally left blank

Appendix

A.1. Numerical Constants

Table A.1 is derived from the Particle Data Group [1], where 10 place accuracy is given for many of the fundamental constants. These numbers are rounded off for ease of numerical calculation, and uncertainties are not shown. See the PDG reference for details. Units in this table are rationalized mks.

A.2. Maxwell's Equations

Table A.2 gives some parameters from Maxwell's equations in different sets of units. Classical mechanics, before relativity, has no scale parameter, and can be described in terms of continuous variables mass, length, and time. Maxwell's equations have a built in scale, the speed of light, and special relativity brings this scale into classical mechanics. Rationalized mks units are widely used in electrical engineering. The potential is measured in volts and the current in amperes — one ampere = one coulomb/sec and volts times amperes = watts. All of these units are household words. 'Rationalized' means that the Maxwell equation $\vec{\nabla} \cdot \vec{E} = \rho/\epsilon_0$ does not contain 4π, which is transferred to Coulomb's law. In cgs units, also called Gaussian units in electromagnetism, the factor of 4π is shifted, so that Coulomb's law is $\vec{F} = q_1 q_2 \hat{r}/r^2$, and the potential energy looks cleaner in the Schrödinger equation. The unit of charge is defined

Table A.1: Fundamental constants from the Particle Data Group [1] in rationalized mks units.

Quantity	Symbol	Value
Speed of light in vacuum	c	2.998×10^{8} m/sec
Elementary charge	e	1.602×10^{-19} coulombs
Electron volt	eV	1.602×10^{-19} joules
Permittivity of free space	ϵ_0	8.85×10^{-12}
Permeability of free space	μ_0	$4\pi \times 10^{-7}$; $\epsilon_0\mu_0 = 1/c^2$
Planck constant	h	4.13×10^{-21} MeV-sec
Conversion constant	hc	1240 eV-nanometers
Planck constant reduced	$\hbar$	6.58×10^{-22} MeV-sec
Conversion constant	$\hbar$c	197.3 eV-nanometers
Conversion constant	$(\hbar c)^2$	0.389 GeV^2 millibarns
Electron mass	m_e	9.11×10^{-31} kg
Electron mass	m_e	0.511 MeV/c^2
Proton mass	m_p	938.3 MeV/c^2
Atomic mass unit	amu ^{12}C/12	931.5 MeV/c^2
W$^{\pm}$ mass; width	M_W; Γ_W	80.4; 2.08 GeV
Z mass; width	M_Z; Γ_Z	91.2; 2.49 GeV
Fine structure constant	$\alpha = e^2/(4\pi\epsilon_0\hbar c)$	1/137.036 dimensionless
Classical electron radius	$r_e = e^2/(4\pi\epsilon_0 m_e c^2)$	2.82×10^{-6} nanometers
Electron Compton wavelength	$\lambda_e/(2\pi) = r_e/\alpha$	3.86×10^{-4} nanometers
Bohr magneton	$\mu_B = e\hbar/(2m_e)$	5.79×10^{-11} MeV/T
Nuclear magneton	$\mu_N = e\hbar/(2m_p)$	3.15×10^{-14} MeV/T
Avogadro's number	N_{Av}	6.02×10^{23} atoms/mole
Boltzmann constant	k	8.62×10^{-5} eV/Kelvin
Strong coupling constant	$\alpha_s(Q^2) = 2\pi/((11 - 2n_f/3)ln(Q/\Lambda))$	$\alpha_s(M_Z^2) = 0.14$ $\Lambda = 0.3$ GeV
Fermi coupling constant	G_F	1.166×10^{-5} GeV^{-2}
Weak mixing angle	$\sin^2\theta_W$	0.231

Table A.2: Parameters from Maxwell's equations in different sets of units.

Quantity	Rationalized mks units	Gaussian units	Natural units
Elementary charge	1.602×10^{-19} coulombs	4.8×10^{-10} esu	0.30
Unit of length	meters	centimeters	1/E
Unit of energy	joules	ergs	MeV
Electrostatic potential	$q_1/(4\pi\epsilon_0 r)$; volts	q_1/r; esu/cm erg/esu	$q_1/(4\pi r)$; E
Electric field $\vec{E} = -\vec{\nabla}V(r)$	$q_1\hat{r}/(4\pi\epsilon_0 r^2)$; volts/m	$q_1\hat{r}/r^2$ esu/cm^2; dyne/esu	$q_1\hat{r}/(4\pi r^2)$; E^2
Electrostatic energy	$(q_1 q_2)/(4\pi\epsilon_0 r)$; joules	$(q_1 q_2)/r$; ergs	$(q_1 q_2)/(4\pi r)$; E
Lorentz force	$q\vec{E} + q\vec{v} \times \vec{B}$; newtons	$q\vec{E} + q(\vec{v}/c) \times \vec{B}$; dyne	$q\vec{E} + q\vec{v} \times \vec{B}$; E^2

so that the force is in dynes. The elementary charge in Gaussian units is the charge in mks units multiplied by 3×10^9, or c/10 in cgs. In Gaussian units the $\vec{E}$ and $\vec{B}$ fields have the same dimensions. This is also true in natural units. This is obvious from the Lorentz force in Table A.2. Another nice feature of Gaussian units is that the electric and magnetic dipole moments have the same dimensions. The electric dipole moment is $\vec{d} = q\vec{a}$, where $\vec{a}$ is the spacial separation of the charge with respect to the center of mass, and the magnetic dipole moment is proportional to $\mu_e = e\hbar/(2mc)$, where $\hbar/(mc)$ is the electron Compton wave length divided by 2π, so it is also a charge times a distance.

A.3. Special Relativity

Repeated indices are summed over unless explicitly stated otherwise. Greek letters $(\mu\nu\rho\sigma)$ refer to four dimensions, and Latin letters (ijk) refer to three dimensions. Four vectors in our metric are real:

$$x^\mu = (x^0, \vec{x}), \quad x_\nu = (x_0, -\vec{x}),$$

$$x^2 = x \cdot x = x^\mu x_\mu = g_{\mu\nu} x^\mu x^\nu = x_0^2 - \vec{x}^2, \quad x^0 = ct,$$

where the metric tensor is

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad g^2 = 1, \quad \mathrm{Tr}(g^2) = 4. \tag{A.1}$$

The square of the distance between two points in four-dimensional space is

$$\begin{gathered}(x-y)^2 = (x^0 - y^0)^2 - (\vec{x} - \vec{y})^2 > 0 \text{ for time-like distance,} \\ \text{and } (x-y)^2 < 0 \ \text{ for space-like distance.}\end{gathered} \tag{A.2}$$

Events separated by a time-like distance can be connected by a light signal. They are time ordered, and a Lorentz frame exists in which $\vec{x} - \vec{y} = 0$. Similarly events separated by a space-like distance cannot be causally connected, and a Lorentz frame exists in which the two events are time coincident: $x^0 = y^0$. This feature is used in evaluating commutators of field operators separated by a space-like distance.

Lorentz transformations are real:

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu}. \tag{A.3}$$

To transform onto a frame moving along the $+z$ direction with velocity v in units of the speed of light, and the Lorentz factor $\gamma = 1/\sqrt{1-v^2}$, the matrix is

$$\Lambda^{\mu}_{\nu} = \begin{pmatrix} \gamma & 0 & 0 & -v\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -v\gamma & 0 & 0 & \gamma \end{pmatrix}. \tag{A.4}$$

Λ is a real, symmetric matrix, with inverse $\Lambda^{-1} = \Lambda(-v)$, simply reverse the sign of the velocity. Space components of a four vector perpendicular to the direction of the transformation are Lorentz invariant, as is the square of the length of a four vector $x^{\mu} x_{\mu}$.

The energy–momentum four vector for an elementary particle is

$$p^{\mu} = (E, \vec{p}), \quad p^{\mu} p_{\mu} = g_{\mu\nu} p^{\mu} p^{\nu} = E^2 - \vec{p}^{\,2} = m^2. \tag{A.5}$$

One of the nice features of our metric is that the squares of time-like vectors are positive. The four dimensional gradient transforms like a lower index vector:

$$\frac{\partial}{\partial x^\mu} = \partial_\mu = \left(\frac{\partial}{\partial t}, \vec{\nabla}\right). \tag{A.6}$$

The passage to quantum mechanics is

$$p_\mu = (E, -\vec{p}) \to i\partial_\mu = \left(i\frac{\partial}{\partial t}, i\vec{\nabla}\right). \tag{A.7}$$

So, the correct sign appears on the gradient: $\vec{p} \to -i\vec{\nabla}$.

A.4. Relativistic Kinematics

A two to two scattering reaction has initial state four vectors p_1 and p_2, and final state four vectors p_3 and p_4. Four momentum is conserved:

$$p_1 + p_2 = p_3 + p_4;$$

$$\text{this is four equations: } E_1 + E_2 = E_3 + E_4, \text{ and } \vec{p}_1 + \vec{p}_2 = \vec{p}_3 + \vec{p}_4. \tag{A.8}$$

Invariant variables called Mandelstam variables are used throughout this text:

$$\begin{gathered} s = (p_1 + p_2)^2, \quad t = (p_1 - p_3)^2, \quad u = (p_1 - p_4)^2, \\ s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2. \end{gathered} \tag{A.9}$$

These variables are the same in any Lorentz frame. In the center of mass $\vec{p}_1' + \vec{p}_2' = 0$, so $s = (E_1' + E_2')^2$, the square of the total energy available. In a storage ring collider the laboratory is the center of mass, and the energy available is twice the energy of either beam. Normally s is time-like, and t and u are both space-like. If $p_2 = (m_2, 0, 0, 0)$, then $s = m_1^2 + m_2^2 + 2m_2E_1$. In the high energy limit $\sqrt{s} \sim \sqrt{2m_2E_1}$, and the energy available increases as the square root of the incident beam energy.

Kinematics of decays are analyzed in the rest frame of the parent, and that is the center of mass. For a two-body decay $M \to m_1 + m_2$ the final state energies are fixed:

$$E_1 = \frac{M^2 + m_1^2 - m_2^2}{2M}, \quad |\vec{p}_1| = |\vec{p}_2| = \frac{\sqrt{\lambda(M^2, m_1^2, m_2^2)}}{2M},$$

where

$$\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2yz - 2xz. \tag{A.10}$$

Three-body decays $M \to m_1 + m_2 + m_3$ are very common, and are more complicated than two-body decays, because there are two independent kinematic variables. In the rest frame of the parent the three final state momenta all lie in a plane. One particle (m_3) can be eliminated using the conservation laws to obtain a formula for the cosine of the angle between the other two. Often the particle eliminated is an undetected neutrino. The angle formula is

$$\begin{aligned} \cos\theta_{12} = (M^2 + m_1^2 + m_2^2 - m_3^2 + 2E_1E_2 \\ - 2M(E_1 + E_2))/(2|\vec{p}_1||\vec{p}_2|). \end{aligned} \tag{A.11}$$

Setting $|\cos\theta_{12}| = 1$ gives the kinematic boundary curve on an E_1E_2 plot. The particles are collinear on the boundary. The minimum energy of any particle is its mass. The maximum energy of m_1 depends on the masses of the other particles. If $m_3 = 0$, then $E_1^{\max} = (M^2 + m_1^2 - m_2^2)/(2M)$, like a two-body decay. On the other hand, if all of the masses are non-zero, then the maximum energy of m_1 occurs when m_2 and m_3 'bind', giving $E_1^{\max} = (M^2 + m_1^2 - (m_2 + m_3)^2)/(2M)$. Two particles cannot simultaneously be at rest. If $E_2 = m_2$, then $E_1 = ((M - m_2)^2 + m_1^2 - m_3^2)/(2(M - m_2))$.

A.5. Dirac δ Functions

These mildly singular functions are very useful. They appear throughout the text. We summarize here a few of their properties.

$$\delta(x) = 0, x \neq 0;$$

$$\text{for } \epsilon \neq 0, \int_{-\epsilon}^{+\epsilon} \delta(x)dx = 1;$$

$$\int_{-\epsilon}^{+\epsilon} f(x)\delta(x)dx = f(0);$$

$$\delta(-x) = \delta(x);$$

$$\delta(ax) = \frac{1}{a}\delta(x);$$

$$\delta(x^2 - a^2) = \frac{1}{2a}(\delta(x-a) + \delta(x+a));$$

$$\text{if } f(x_0) = 0 \text{ has only one root, then}$$

$$\int \delta(f(x))dx = \int \frac{dx}{df}\delta(f)df = \frac{dx}{df}(x = x_0).$$

For $\theta(x)$, a step function: $\theta(x) = 0, x < 0$, $\theta(x) = 1, x > 0$; $\delta(x) = d\theta/dx$; $\int_{-\epsilon}^{+\epsilon}(d\theta/dx)dx = \theta(+\epsilon) - \theta(-\epsilon) = 1$. The integral representation occurs frequently; in three dimensions:

$$\delta(\vec{x}) = \frac{1}{(2\pi)^3}\int_{-\infty}^{+\infty} d^3k e^{i\vec{k}\cdot\vec{x}}. \tag{A.12}$$

Some continuous functions that become δ functions as a limiting case are discussed in the text.

A.6. Phase Space

A phase space factor appears in the transition rate formula, and in the cross-section formula — the transition rate divided by the flux. It is kinematics, borrowed from a theorem of statistical mechanics that states "consistent with the constraints equal volumes of phase space have equal *a priori* probabilities". The four momentum conserving δ functions are the constraints, and the differentials $d^3p/2E$ are Lorentz invariant volume elements in momentum space. Two bodies are sufficient to integrate the four δ functions. More than two bodies can be partially treated with two-body phase space by defining the appropriate initial state four momentum:$X^\mu = p_0^\mu - \sum_3^n p_j^\mu$. Then the two-body phase space is defined as

$$d_2(PS) = \delta^4(X - p_1 - p_2)\frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2}, \tag{A.13}$$

where X is the four momentum available to final state particles m_1 and m_2. Integration over d^3p_2 in the center of mass of the two bodies, where $X^0 = \sqrt{s}$, and $\vec{X} = 0$ gives:

$$\int d_2(PS) = \delta\left(\sqrt{s} - E_1 - \sqrt{m_2^2 + E_1^2 - m_1^2}\right) \frac{p_1^2 dp_1 d\Omega}{4E_1E_2}. \qquad \text{(A.14)}$$

The argument of the δ function is $f(E_1) = \sqrt{s} - E_1 - \sqrt{m_2^2 + E_1^2 - m_1^2}$, and $dE_1/df = E_2/\sqrt{s}$. Using $p_1dp_1 = E_1dE_1$ we have:

$$\int\int d_2(PS) = \frac{p_1 d\Omega}{4\sqrt{s}} = \frac{\pi\sqrt{\lambda(s, m_1^2, m_2^2)}}{2s} \frac{d\Omega}{4\pi}, \qquad \text{(A.15)}$$

which is a constant times the differential solid angle in the m_1m_2 center of mass. Sometimes the matrix element contains simple functions of the momenta. Examples have been encountered in the text. A simple form was used for two final state neutrinos, where $m_1^2 = m_2^2 = 0$. The general form for non-zero masses, adopted from [2, Appendix B] is

$$\begin{aligned}\int d_2(PS)p_1^\mu p_2^\nu = {}& \frac{\pi}{24}\sqrt{\lambda(s, m_1^2, m_2^2)} \\ & \times \Big\{g^{\mu\nu}\lambda(s, m_1^2, m_2^2)/s^2 \\ & + 2\frac{X^\mu X^\nu}{s^3}(s^2 + s(m_1^2 + m_2^2) - 2(m_1^2 - m_2^2)^2)\Big\}.\end{aligned} \qquad \text{(A.16)}$$

The proof of this formula rests on the claim that, because p_1 and p_2 are integrated out, the general form must be:

$$\int d_2(PS)p_1^\mu p_2^\nu = Ag^{\mu\nu}X^2 + BX^\mu X^\nu, \qquad \text{(A.17)}$$

where A and B are functions only of s, m_1^2, m_2^2. Multiply by $X_\mu X_\nu$:

$$\int d_2(PS)E_1E_2s = \frac{\pi\sqrt{\lambda(s, m_1^2, m_2^2)}}{2}E_1E_2 = (A + B)X^4. \qquad \text{(A.18)}$$

Now multiply by $g_{\mu\nu}$:

$$\int d_2(PS)(E_1E_2 + p^2) = \frac{\pi\sqrt{\lambda(s, m_1^2, m_2^2)}}{2s}\left(E_1E_2 + \frac{\lambda}{4s}\right)$$
$$= (4A + B)X^2. \tag{A.19}$$

We have two equations in two unknowns, A and B:

$$A = \frac{\pi(\lambda(s, m_1^2, m_2^2))^{3/2}}{24s^3}, \quad \text{and}$$
$$B = \frac{\pi\sqrt{\lambda(s, m_1^2, m_2^2)}}{12s^3}(s^2 + s(m_1^2 + m_2^2) - 2(m_1^2 - m_2^2)^2). \tag{A.20}$$

Substituting Equations (A.20) into Equation (A.17) gives Equation (A.16). We have used Equations (A.10) to express E_1 and E_2 in terms of invariants.

As an example of the split of phase space into two bodies plus everyone else, let us consider three bodies both ways: the usual way that involves dE_1dE_2, and using Equation (A.15) for the 2–3 phase space evaluated in the 2–3 center of mass. For the decay $M \to m_1 + m_2 + m_3$ we define:

$$d_3(PS) = \delta(M - E_1 - E_2 - E_3)\delta(\vec{p}_1 + \vec{p}_2 + \vec{p}_3)\frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2}\frac{d^3p_3}{2E_3}. \tag{A.21}$$

This expression has been evaluated in the text, Equation (2.86):

$$d_3(PS) = \pi^2 dE_1dE_2, \quad \text{so}$$
$$\int d_3(PS) = \pi^2 \int_{m_1}^{E_1^{\max}} \int_{E_2^{\min}}^{E_2^{\max}} dE_1dE_2. \tag{A.22}$$

The limits on E_2 come from the boundary curve Equation (A.11). Square both sides with $|\cos\theta_{12}| = 1$, and solve the quadratic equation (2.69) to obtain

$$\int_{E_2^{\min}}^{E_2^{\max}} dE_2 = \frac{p_1\sqrt{\lambda(s, m_2^2, m_3^2)}}{s},$$
$$d_3(PS) = \pi^2 \int_{m_1}^{E_1^{\max}} dE_1 \frac{p_1\sqrt{\lambda(s, m_2^2, m_3^2)}}{s},$$

where

$$s = M^2 + m_1^2 - 2ME_1. \tag{A.23}$$

The significance of the definition of s will be clear presently. Now go into the rest frame of m_2, m_3, and evaluate the two body phase space as we have done in Equation (A.15), and then tack on the extra factor for m_1. This procedure will not give $dE_1 dE_2$, because m_2 has already been integrated out in $d_2(PS)$, but it will give the correct form for $\int dE_1$. So how do we get to the m_2, m_3 rest frame? In the rest frame of the parent M particles 2 and 3 balance the momentum of particle 1, so the total momentum of interest is $\vec{p}_2 + \vec{p}_3 = -\vec{p}_1$, and the total energy is $E_2 + E_3$, so the Lorentz transformation is

$$\vec{v} = -\frac{\vec{p}_1}{E_2 + E_3}, \quad \gamma = \frac{E_2 + E_3}{\sqrt{M^2 + m_1^2 - 2ME_1}},$$

and the total energy available in the center of mass of m_2, m_3 is

$$\sqrt{s} = \gamma(E_2 + E_3 - vp_1) = \sqrt{M^2 + m_1^2 - 2ME_1}. \tag{A.24}$$

This explains the definition of s in Equation (A.23). As a check, if $E_1 = E_1^{\max}, s = (m_2 + m_3)^2$, so there is no extra energy available in the m_2, m_3 center of mass. For the phase space we need the momentum in the m_2, m_3 center of mass, our usual formula:

$$p_{23}^* = \frac{\sqrt{\lambda(s, m_2^2, m_3^2)}}{2\sqrt{s}}. \tag{A.25}$$

The two-body phase space factor is

$$\int d_2(PS) = \pi \frac{\sqrt{\lambda(s, m_2^2, m_3^2)}}{2s},$$

and the three-body phase space is

$$d_3(PS) = \frac{d^3 p_1}{2E_1} \pi \frac{\sqrt{\lambda(s, m_2^2, m_3^2)}}{2s};$$

$$\int d_3(PS) = \pi^2 \int_{m_1}^{E_1^{\max}} dE_1 \frac{p_1 \sqrt{\lambda(s, m_2^2, m_3^2)}}{s}. \tag{A.26}$$

We have derived Equation (A.23) in two ways. The integral is not simple, since s depends on E_1.

A.7. Cross-Section and Decay Rate Formulas

The cross-section and decay rate formulas are in relativistically invariant form. They involve the transition matrix element $|\langle f|\,T\,|i\rangle|^2$ that may be calculated from the Feynman rules that are summarized in this Appendix. The cross section for a two-to-two scattering is

$$d\sigma_{i\to f} = (2\pi)^4\delta^4(p_1+p_2-p_3-p_4)\frac{d^3p_3}{(2\pi)^3 2E_3}\frac{d^3p_4}{(2\pi)^3 2E_4} \times |\langle f|\,T\,|i\rangle|^2 \frac{1}{4(E_2|\vec{p}_1|+E_1|\vec{p}_2|)}. \tag{A.27}$$

For more final state particles add their four momenta into the δ function and tack on extra phase space factors $d^3p/((2\pi)^3 2E)$. The last part is the flux. This formula can be simplified in the center of mass if all of the masses can be neglected, i.e. $s \gg m^2$ and all energies are equal:

$$\frac{d\sigma}{d\Omega} = \frac{|\langle f|\,T\,|i\rangle|^2}{64\pi^2 s}. \tag{A.28}$$

The T matrix element is dimensionless, and $1/s$ in GeV^{-2} can be converted directly into millibarns using the conversion constant in Table A.1: $1\,\mathrm{GeV}^{-2} = 0.389\,\mathrm{mb}$.

The decay rate formula is for $M \to m_1 + m_2 + m_3 + \cdots$ in the M rest frame is:

$$d\Gamma_{i\to f} = \frac{(2\pi)^4}{2M}\delta(M-E_1-E_2-E_3-\cdots)\delta(\vec{p}_1+\vec{p}_2+\vec{p}_3+\cdots) \times |\langle f|\,T\,|i\rangle|^2 \prod_{i=1}^{n}\frac{d^3p_i}{(2\pi)^3 2E_i}. \tag{A.29}$$

Γ has energy units. The T matrix element is dimensionless for a three-body final state. The δ function is E^{-4}, phase space is E^6, and $1/M$ gives Γ in MeV. To convert to s^{-1}, divide by $\hbar = 6.58 \times 10^{-22}$ MeV-s.

A.8. Dirac Electron Theory

Dirac's linear equation is

$$(\vec{\alpha} \cdot \vec{p} + \beta m)\psi(\vec{x}, t) = i\frac{\partial \psi}{\partial t}, \quad \text{where}$$

$$\alpha_i \alpha_j + \alpha_j \alpha_i = 2\delta_{ij}, \quad \alpha_i \beta + \beta \alpha_i = 0, \quad \beta^2 = 1.$$

In derivative form:

$$(-i\vec{\alpha} \cdot \vec{\nabla} + \beta m)\psi = i\frac{\partial \psi}{\partial t}. \tag{A.30}$$

The matrices are

$$\alpha_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix},$$

$$\alpha_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{A.31}$$

$\vec{\alpha}$ and β are hermitian matrices: $\vec{\alpha}^\dagger = \vec{\alpha}$, $\beta^\dagger = \beta$. The relativistic form is

$$\vec{\gamma} = \beta\vec{\alpha}, \quad \gamma^0 = \beta, \quad i\gamma^\mu \frac{\partial \psi}{\partial x^\mu} = m\psi. \tag{A.32}$$

The γ matrices satisfy:

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}, \quad \text{and} \quad \gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0. \tag{A.33}$$

The adjoint equation is

$$-i\frac{\partial\bar{\psi}}{\partial x^\mu}\gamma^\mu = m\bar{\psi}, \quad \bar{\psi} = \psi^\dagger\gamma^0,$$

$$\text{and} \quad \frac{\partial\bar{\psi}\gamma^\mu\psi}{\partial x^\mu} = 0 \text{ is the conserved current.} \tag{A.34}$$

$\bar{\psi}\gamma^\mu\psi$ is a four-vector under Lorentz transformations. Spinors come from plane wave solutions:

$$\psi(\vec{x},t) = u^{(s)}(p)e^{-ip_\mu x^\mu}, \quad \frac{\partial\psi}{\partial x^\mu} = -ip_\mu\psi, \quad \gamma^\mu p_\mu\psi = m\psi,$$

$$(\not{p} - m)u^{(s)}(p) = 0, \quad \bar{u}^{(s)}(p)(\not{p} - m) = 0. \tag{A.35}$$

The spin index $s = 1, 2$ for electrons and $s = 1, 2$ for positrons given below, based on electron spinors with indices $s = 4, 3$. An important operator for the weak interactions that anticommutes with all four γ matrices is γ^5:

$$\begin{aligned}
\gamma^5 &= i\gamma^0\gamma^1\gamma^2\gamma^3 = \frac{i}{4!}\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma, \\
\gamma^{5\dagger} &= \gamma^5, \\
(\gamma^5)^2 &= 1, \quad \gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0, \\
\gamma^5 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},
\end{aligned} \tag{A.36}$$

$$\begin{aligned}
\epsilon_{0123} &= -\epsilon^{0123} = +1, \quad \epsilon_{0213} = -1, \\
&\qquad \text{interchange any two indices flips the sign,} \\
\epsilon_{\mu\mu\rho\sigma} &= 0, \quad \text{any two indices the same.} \\
-\epsilon^{\alpha\beta\mu\nu}\epsilon_{\alpha\beta\sigma\tau} &= 2(\delta^\mu_\sigma\delta^\nu_\tau - \delta^\mu_\tau\delta^\nu_\sigma), \\
-\epsilon^{\alpha\beta\gamma\nu}\epsilon_{\alpha\beta\gamma\tau} &= 6\delta^\nu_\tau, \\
-\epsilon^{\alpha\beta\gamma\delta}\epsilon_{\alpha\beta\gamma\delta} &= 4! = 24.
\end{aligned} \tag{A.37}$$

The explicit forms for the electron spinors are

$$u^{(1)}(p) = \sqrt{E+m}\begin{pmatrix} 1 \\ 0 \\ p_z/(E+m) \\ (p_x+ip_y)/(E+m) \end{pmatrix},$$

$$u^{(2)}(p) = \sqrt{E+m}\begin{pmatrix} 0 \\ 1 \\ (p_x-ip_y)/(E+m) \\ -p_z/(E+m) \end{pmatrix}, \tag{A.38}$$

and the positron spinors are

$$v^{(1)}(p) = \sqrt{E+m}\begin{pmatrix} (p_x-ip_y)/(E+m) \\ -p_z/(E+m) \\ 0 \\ 1 \end{pmatrix},$$

$$v^{(2)}(p) = \sqrt{E+m}\begin{pmatrix} p_z/(E+m) \\ (p_x+ip_y)/(E+m) \\ 1 \\ 0 \end{pmatrix}. \tag{A.39}$$

Energies are positive for all spinors, and the electron and positron masses are the same. The positron spinors obey the Dirac equation:

$$(\not{p}+m)v^s(p) = 0. \tag{A.40}$$

The spinors are normalized:

$$\begin{aligned} u^{(s)\dagger}(p)u^{s'}(p) &= v^{(s)\dagger}(p)v^{s'}(p) = 2E\delta_{ss'}, \\ \bar{u}^s(p)u^{s'}(p) &= -\bar{v}^s(p)v^{s'}(p) = 2m\delta_{ss'}. \end{aligned} \tag{A.41}$$

The completeness relations are

$$\begin{aligned} \sum_{s=1,2} u^{(s)}(p)\bar{u}^{(s)}(p) &= \not{p}+m, \\ \sum_{s=1,2} v^{(s)}(p)\bar{v}^{(s)}(p) &= \not{p}-m. \end{aligned} \tag{A.42}$$

A.9. The Feynman Rules

Rules for drawing a Feynman diagram and constructing from it a probability amplitude for the calculation of such physical observables as cross sections and decay rates were first developed for the quantum electrodynamics of electrons, positrons, and photons. Spin 1/2 quarks carry charge, and are assumed to obey the Dirac equation, so QED applies directly to the quarks as well. These rules have been extended to charged and neutral current weak interactions of leptons, quarks, $W^{\pm}$ and Z bosons. The EWK theory, based on vector-axial vector coupling of the spin 1/2 particles, combines weak and electromagnetic interactions. Feynman rules have also been developed for the Higgs boson, where the coupling to spin 1/2 and spin 1 particles is scalar. And last but not least there are Feynman rules for QCD of quarks and gluons, a vector non-abelian gauge theory.

The rules are illustrated in a series of diagrams. In each case the T matrix element $-i \langle f | T | i \rangle$ is given by the product of the spinors, vertex factors, and propagators. Figure A.1 show the rules for quantum electrodynamics, reproduced from Figure 4.2 of Chapter 4. The coupling constant is the elementary charge multiplied by Q^f. For the electron $Q^f = -1$. Algebraic signs in the amplitudes are important, as demonstrated in Chapter 9, where we calculated the effect on the positronium hyperfine splitting of the annihilation diagram. Figures A.2 and A.3 give the rules for charged current weak interactions. Neutral current weak interactions of fermions are summarized in Figure A.4. For this figure one needs Table A.3, where $\sin^2\theta_W = 0.23$, based on the definitions:

$$g_V^f = T_3^f - 2Q^f \sin^2\theta_W, \quad g_A^f = T_3^f. \tag{A.43}$$

The neutrino coupling to the Z is the same as the coupling of all charged and neutral leptons to the W, but the coupling of charged fermions is charge dependent. The experimental value of $\sin^2\theta_W$ is near 0.25, and that leads to a very small vector coupling of the charged leptons to the Z. Z calculations are often simpler using the orthogonal helicity states $g_L = (g_V + g_A)/2$ and $g_R = (g_V - g_A)/2$. Figure A.5 gives the rules for single vector boson production or decay. Outgoing W would be ϵ_μ^* and incoming ϵ_μ, just like the photon,

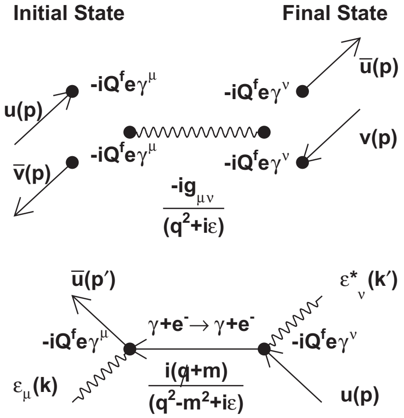

Initial State
Final State
u(p)
$\bar{\mathrm{v}}$(p)
-iQ^{f}eγ$^\mu$
-iQ^{f}eγ$^\mu$
-iQ^{f}eγ$^\nu$
-iQ^{f}eγ$^\nu$
$\bar{\mathrm{u}}$(p)
v(p)
-ig$_{\mu\nu}$ / (q^2+iε)
$\bar{\mathrm{u}}$(p′)
ε*$_\nu$(k′)
γ+e$^-$→ γ+e$^-$
-iQ^{f}eγ$^\mu$
-iQ^{f}eγ$^\nu$
ε$_\mu$(k)
i(q̸+m) / (q^2-m^2+iε)
u(p)

Figure A.1: Feynman rules for charged fermions — e's, μ's, τ's and quarks — and photons. Q^f is the fermion charge, and e is positive. For an electron $Q^f = -1$, and for a u quark $Q^f = +2/3$. The correct four momentum must be inserted in the spinors. Antifermion lines are directed opposite to the momentum. Real photons have momentum k, and the four momentum in the propagators is $q = p_{\text{in}} - p_{\text{out}}$. Begin with the left vertex, sandwich $\bar{u}(p_3)(-ie\gamma^\mu)u(p_1)$. Couple the other vertex with the photon propagator $-ig_{\mu\nu}/q^2$. The sandwich with three $-i$ gives $-i\langle f|\,T\,|i\rangle$. One of the diagrams for Compton scattering is shown on the bottom. The electron propagator involves $q = k - p'$. Interchanging two electron lines or exchanging an outgoing positron for an incoming electron changes the sign of the amplitude. Interchanging two photon lines gives a plus sign.

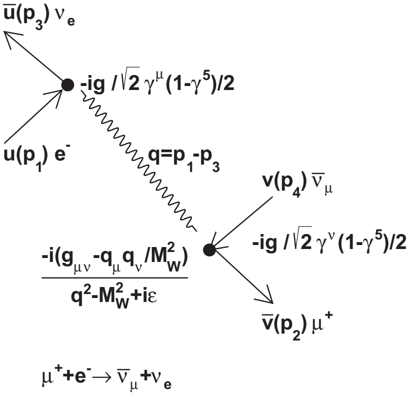

$\bar{\mathrm{u}}$(p$_3$) ν$_e$
-ig /√2 γ$^\mu$(1-γ^5)/2
u(p$_1$) e$^-$
q=p$_1$-p$_3$
v(p$_4$) $\bar{\nu}_\mu$
-ig /√2 γ$^\nu$(1-γ^5)/2
-i(g$_{\mu\nu}$-q$_\mu$q$_\nu$/M$_W^2$) / q^2-M$_W^2$+iε
$\bar{\mathrm{v}}$(p$_2$) μ$^+$
μ$^+$+e$^-$→ $\bar{\nu}_\mu$+ν$_e$

Figure A.2: Feynman rules for charged current weak interactions of charged leptons, e's, μ's, and τ's, and their neutrinos coupled to charged weak $W^\pm$ bosons. The constant $g = 2M_W(\sqrt{2}G_F)^{1/2}$, and $g\sin\theta_W = e$, the elementary charge. A particular reaction, the annihilation of muonium, a t-channel process that resembles muon decay, is shown. See Problem 9.6.

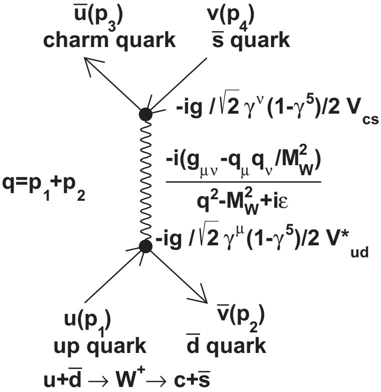

$\bar{u}(p_3)$
charm quark
$v(p_4)$
$\bar{s}$ quark
$-ig/\sqrt{2}\,\gamma^\nu(1-\gamma^5)/2\,V_{cs}$
$q=p_1+p_2$
$\frac{-i(g_{\mu\nu}-q_\mu q_\nu/M_W^2)}{q^2-M_W^2+i\varepsilon}$
$-ig/\sqrt{2}\,\gamma^\mu(1-\gamma^5)/2\,V^*_{ud}$
$u(p_1)$
up quark
$\bar{v}(p_2)$
$\bar{d}$ quark
$u+\bar{d}\rightarrow W^+\rightarrow c+\bar{s}$

Figure A.3: Feynman rules for charged current weak interactions of quarks. The specific process shown is the dominant s-channel parton model reaction for W^+ production: $u+\bar{d}\rightarrow W^+\rightarrow c+\bar{s}$. The CKM matrix element V_{ud} is written $\cos\theta_c$ in the text before the full 3×3 matrix is introduced. The convention chosen for the CKM matrix elements is described in the text; V for outgoing u-type quarks, and V^* for outgoing d-type quarks. For production of a real on mass W the propagator is replaced by a relativistic Breit–Wigner $\epsilon\rightarrow\Gamma_W M_W$, and the coupling to the various final states of W decay give the branching ratios. 100% of the W's decay with lifetime 3×10^{-25} s.

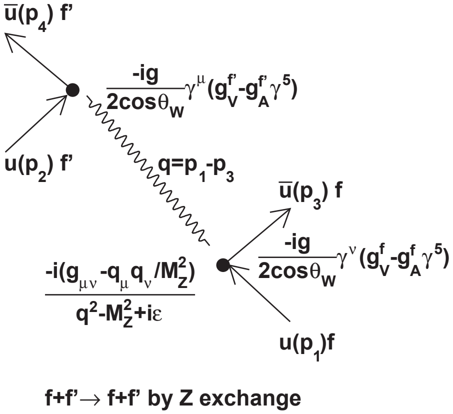

$\bar{u}(p_4)$ f'
$\frac{-ig}{2\cos\theta_W}\gamma^\mu(g_V^{f'}-g_A^{f'}\gamma^5)$
$u(p_2)$ f'
$q=p_1-p_3$
$\bar{u}(p_3)$ f
$\frac{-ig}{2\cos\theta_W}\gamma^\nu(g_V^f-g_A^f\gamma^5)$
$\frac{-i(g_{\mu\nu}-q_\mu q_\nu/M_Z^2)}{q^2-M_Z^2+i\varepsilon}$
$u(p_1)$f
f+f'$\rightarrow$ f+f' by Z exchange

Figure A.4: Feynman rules for weak neutral current coupling to fermions. The coupling depends on the weak isospin, the charge, and the angle θ_W. Values are summarized in Table A.3.

Table A.3: The coupling of weak neutral current to fermions.

Fermion	g_V	g_A
ν_e, ν_μ, ν_τ	$+1/2$	$+1/2$
e^-, μ^-, τ^-	$-1/2 + 2\sin^2\theta_W$	$-1/2$
u, c, t	$+1/2 - 4\sin^2\theta_W/3$	$+1/2$
d, s, b	$-1/2 + 2\sin^2\theta_W/3$	$-1/2$

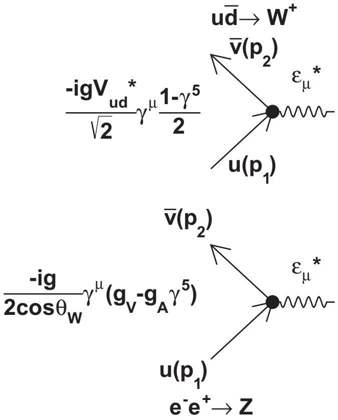

$u\bar{d} \to W^+$
$\bar{v}(p_2)$
$\varepsilon_\mu{}^*$
$\frac{-igV_{ud}{}^*}{\sqrt{2}}\gamma^\mu\frac{1-\gamma^5}{2}$
$u(p_1)$
$\bar{v}(p_2)$
$\varepsilon_\mu{}^*$
$\frac{-ig}{2\cos\theta_W}\gamma^\mu(g_V-g_A\gamma^5)$
$u(p_1)$
$e^-e^+ \to Z$

Figure A.5: Feynman rules for production of real vector bosons $e^-e^+ \to Z$ and $u\bar{d} \to W^+$. The outgoing vector bosons are represented by their polarization vectors, just like the photon. The coupling constant $g = 2M_W(\sqrt{2}G_F)^{1/2}$, and $g\sin\theta_W = e$, the elementary charge.

although the sum over polarizations has the extra term:

$$\sum_\lambda \epsilon_\mu^{\lambda *}\epsilon_\nu^\lambda = -g_{\mu\nu} + \frac{p_\mu p_\nu}{M_W^2}, \tag{A.44}$$

where p_μ is the W four momentum.

Figure A.6 summarizes the Higgs couplings to itself, to the fermions, and to the vector bosons. These mass dependent couplings are all proportional to $1/v$, the vacuum expectation value of the scalar field that breaks the symmetry and gives masses to $W^\pm$ and Z.

The scalar field Lagrangian in the model is

$$\mathcal{L} = D_\mu \phi^\dagger D^\mu \phi + \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2,$$
$$v = \frac{\mu}{\sqrt{\lambda}} = \frac{1}{(\sqrt{2}G_F)^{1/2}}. \tag{A.45}$$

The covariant derivatives generate the masses for the vector fields W and Z. The fermions have to be treated separately in the model. This is all described in Chapter 8, where these couplings are derived. As shown in Figure A.6, the scalar coupling does not involve the γ matrices. The metric tensor in the coupling to the vector bosons takes care of the polarization four vectors. The $W^\pm$ or Z can radiate a Higgs, but a Higgs cannot radiate a Z. The Higgs analysis in Chapter 8 is a good demonstration of the fact that the Feynman rules can be obtained by inspection of the covariant derivatives in the Lagrangian.

We end the section on the Feynman Rules with a brief discussion of quark–gluon coupling in QCD. The goal is to obtain the color factor 4/3 that was used in Chapter 9 in discussing the non-relativistic quark–antiquark potential. Figure A.7 shows the

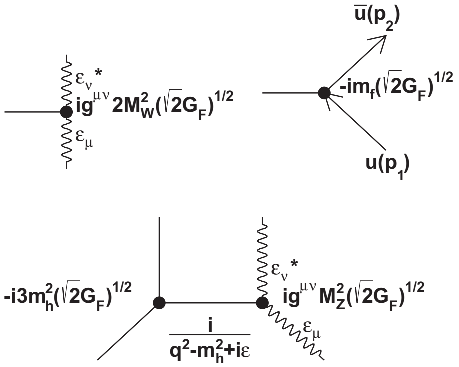

$\varepsilon_\nu{}^*$
$ig^{\mu\nu}2M_W^2(\sqrt{2}G_F)^{1/2}$
ε_μ
$\bar{u}(p_2)$
$-im_f(\sqrt{2}G_F)^{1/2}$
$u(p_1)$
$-i3m_h^2(\sqrt{2}G_F)^{1/2}$
$\frac{i}{q^2-m_h^2+i\varepsilon}$
$\varepsilon_\nu{}^*$
$ig^{\mu\nu}M_Z^2(\sqrt{2}G_F)^{1/2}$
ε_μ

Figure A.6: Feynman rules for Higgs self coupling and coupling to $W^\pm, Z$, and fermions. The Higgs propagator is also shown in the lower diagram. External Higgs lines multiply the matrix element by one. Note the factor of two in the W coupling compared to the Z, and that the boson couplings are proportional to M^2, while the fermions are proportional to m. All of the Higgs formulas involve the vacuum expectation value of the scalar field v, which has been replaced via $v = (\sqrt{2}G_F)^{-1/2}$.

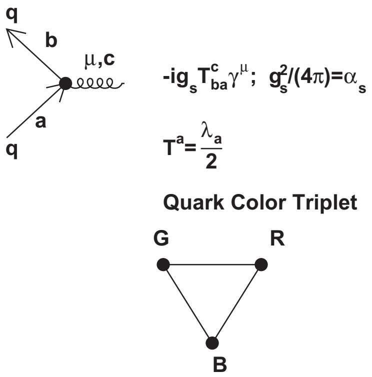


Figure A.7: Feynman Rule for quark–gluon coupling. A quark of any flavor (u, d, c, s, t, b) and color index a (three values: red, green, and blue) emits a gluon with index c (1–8), and becomes a quark of the same flavor and color index b. The eight 3×3 operators $T^c = \lambda_c/2$ are given in this Appendix. The vector coupling index is μ. The coupling constant $g_s = \sqrt{4\pi\alpha_s}$.

quark–gluon vertex. There are eight gluons, labeled with the eight operators of the SU(3) color group. Two of the matrices are diagonal. The usual representation of the hermitian λ_a 3×3 matrices, that operate in the triplet space of 3 and 3* is similar to the Pauli matrices, and is attributed to Gell-Mann:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$\lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{A.46}$$

The red, green, and blue color state vectors of the triplet shown in Figure A.7 are

$$|R\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |G\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |B\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{A.47}$$

Linear combinations of the λ's are the raising and lowering operators of the triplet. For example,

$$\begin{aligned} &\frac{\lambda_1 + i\lambda_2}{2}|G\rangle = |R\rangle, \quad \frac{\lambda_4 + i\lambda_5}{2}|B\rangle = |R\rangle, \quad \text{and} \\ &\frac{\lambda_6 + i\lambda_7}{2}|B\rangle = |G\rangle. \end{aligned} \tag{A.48}$$

In order to obtain the color factor for charmonium or bottomonium we need to average over all initial colors R, G, and B, and sum the squares of all of the contributing gluons. While there is group theoretical machinery to do this, we may as well simply work it out with the matrices and the column vectors. There are only 24 terms, and seven are zero. Thus,

$$\begin{aligned} \lambda_1|B\rangle &= \lambda_2|B\rangle = \lambda_3|B\rangle = \lambda_4|G\rangle \\ &= \lambda_5|G\rangle = \lambda_6|R\rangle = \lambda_7|R\rangle = 0. \end{aligned} \tag{A.49}$$

$$\begin{aligned} &\lambda_1|R\rangle = |G\rangle, \quad \lambda_2|R\rangle = i|G\rangle, \quad \lambda_3|R\rangle = |R\rangle, \quad \lambda_1|G\rangle = |R\rangle, \\ &\lambda_2|G\rangle = -i|R\rangle, \quad \lambda_3|G\rangle = -|G\rangle. \\ &\lambda_4|R\rangle = |B\rangle, \quad \lambda_5|R\rangle = i|B\rangle, \quad \lambda_6|G\rangle = |B\rangle, \\ &\lambda_4|B\rangle = |R\rangle, \quad \lambda_5|B\rangle = -i|R\rangle, \quad \lambda_6|B\rangle = |G\rangle. \\ &\lambda_7|G\rangle = i|B\rangle, \quad \lambda_7|B\rangle = -i|G\rangle, \quad \lambda_8|R\rangle = |R\rangle/\sqrt{3}, \\ &\lambda_8|G\rangle = |G\rangle/\sqrt{3}, \quad \lambda_8|B\rangle = -2|B\rangle/\sqrt{3}. \end{aligned} \tag{A.50}$$

So in the diagram, starting with a blue quark, we have:

$$\begin{aligned} &|B\rangle \to \lambda_8 + |B\rangle = -2/\sqrt{3}, \quad |B\rangle \to \lambda_6 + |G\rangle = 1, \\ &|B\rangle \to \lambda_7 + |G\rangle = -i, \\ &|B\rangle \to \lambda_4 + |R\rangle = 1, \quad |B\rangle \to \lambda_5 + |R\rangle = -i. \end{aligned} \tag{A.51}$$

The sum of the magnitude squared is $|\lambda_8|^2 + |\lambda_6|^2 + |\lambda_7|^2 + |\lambda_4|^2 + |\lambda_5|^2 = 4/3 + 1 + 1 + 1 + 1 = 16/3$. Divide by four because $T = \lambda/2$ to get 4/3. Each initial color contributes the same, and the result is divided by three for a colorless $q\bar{q}$ state. This gives the factor in Equation (9.40). It appears that we have ignored the other vertex — the one with the anti-colors. This is OK because the gluon transfers the color information from one vertex to the other, and the generators for the 3^* are

$$\lambda'_a = -\lambda^*_a = -\tilde{\lambda}_a, \tag{A.52}$$

where the tilde means the transpose of the hermitian matrix. The same gluon couples an $R\bar{R}$ initial state to a $G\bar{G}$ final state.

A.10. Trace Theorems

After constructing the amplitude from the Feynman rules, it is necessary to multiply it by its complex conjugate, and to sum over all spins. To be specific, we consider one of the amplitudes for Möller scattering:

$$\begin{aligned} -i\langle f|\,T\,|i\rangle &= (ie\bar{u}(p_3)\gamma^\mu u(p_1))\left(\frac{-ig_{\mu\nu}}{(p_1-p_3)^2}\right)(ie\bar{u}(p_4)\gamma^\nu u(p_2)), \\ \langle f|\,T\,|i\rangle &= -\frac{e^2}{q^2}\bar{u}(p_3)\gamma^\mu u(p_1)\bar{u}(p_4)\gamma_\mu u(p_2). \end{aligned} \tag{A.53}$$

Now insert spin indices, multiply by the complex conjugate, and sum over all spins:

$$\begin{aligned} \frac{1}{4}\sum_{\text{spins}} |\langle f|\,T\,|i\rangle|^2 = \frac{e^4}{4q^4} &\sum_{s,s',s'',s'''} \bar{u}^{s''}(p_3)\gamma^\mu u^s(p_1)\bar{u}^{s'''}(p_4)\gamma_\mu u^{s'}(p_2) \\ &\times \bar{u}^{s'}(p_2)\gamma_\lambda u^{s'''}(p_4)\bar{u}^s(p_1)\gamma^\lambda u^{s''}(p_3). \end{aligned} \tag{A.54}$$

The photon propagator contains the sum over intermediate photon spin states. The factor of 1/4 averages over the initial spins. The sum over s' can be replaced by the completeness theorem. Then write the expression in component form:

$$\frac{1}{4}\sum_{\text{spins}}|\langle T\rangle|^2 = \frac{e^4}{4q^4}\sum_{s,s'',s'''}\bar{u}^{s''}_{\alpha}(\gamma^{\mu})_{\alpha\beta}u^{s}_{\beta}\bar{u}^{s'''}_{\gamma}(\gamma_{\mu})_{\gamma\nu} \\ \times(\not{p}_2+m)_{\nu\tau}(\gamma_{\lambda})_{\tau\rho}u^{s'''}_{\rho}\bar{u}^{s}_{\delta}(\gamma^{\lambda})_{\delta\sigma}u^{s''}_{\sigma}. \quad \text{(A.55)}$$

In component form you may shift objects around. Shuffle the spinors to obtain

$$\frac{e^4}{4q^4}(\not{p}_3+m)_{\sigma\alpha}(\gamma^{\mu})_{\alpha\beta}(\not{p}_1+m)_{\beta\delta}(\gamma^{\lambda})_{\delta\sigma} \\ \times(\not{p}_4+m)_{\rho\gamma}(\gamma_{\mu})_{\gamma\nu}(\not{p}_2+m)_{\nu\tau}(\gamma_{\lambda})_{\tau\rho} \\ = \frac{e^4}{4q^4}\mathrm{Tr}((\not{p}_3+m)\gamma^{\mu}(\not{p}_1+m)\gamma^{\lambda}) \\ \times\mathrm{Tr}((\not{p}_4+m)\gamma_{\mu}(\not{p}_2+m)\gamma_{\lambda}). \quad \text{(A.56)}$$

We collect the trace theorems from Chapter 4. A few relations that are not traces are the following:

$$\gamma^{\mu}\gamma_{\mu} = (\gamma^0)^2-(\vec{\gamma})^2 = 4, \quad \text{(A.57)}$$

$$\gamma^{\mu}\not{a}\gamma_{\mu} = -2\not{a}, \quad \text{(A.58)}$$

$$\gamma^{\mu}\not{a}\not{b}\gamma_{\mu} = 4a\cdot b, \quad \text{(A.59)}$$

$$\gamma^{\mu}\not{a}\not{b}\not{c}\gamma_{\mu} = -2\not{c}\not{b}\not{a}. \quad \text{(A.60)}$$

For trace calculations, the trace of an odd number of γ matrices vanishes. For two, we have:

$$\mathrm{Tr}(\not{a}\not{b}) = 4a\cdot b, \quad \text{or} \quad \mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}) = 4g^{\mu\nu}. \quad \text{(A.61)}$$

For the product of four gamma matrices, the result is

$$\mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}\gamma^{\rho}\gamma^{\sigma}) = 4(g^{\mu\nu}g^{\rho\sigma}+g^{\mu\sigma}g^{\nu\rho}-g^{\mu\rho}g^{\nu\sigma}). \quad \text{(A.62)}$$

The order can be reversed in Equation (A.62), and this inversion is possible in the trace of a larger product of an even number of γ matrices. The trace of a product of six γ matrices can be reduced to five traces of four γ matrices as follows:

$$\begin{aligned}\mathrm{Tr}(\not{a}\not{b}\not{c}\not{d}\not{e}\not{f}) &= (a\cdot b)\mathrm{Tr}(\not{c}\not{d}\not{e}\not{f}) - (a\cdot c)\mathrm{Tr}(\not{b}\not{d}\not{e}\not{f})\\ &\quad + (a\cdot d)\mathrm{Tr}(\not{b}\not{c}\not{e}\not{f}) - (a\cdot e)\mathrm{Tr}(\not{b}\not{c}\not{d}\not{f})\\ &\quad + (a\cdot f)\mathrm{Tr}(\not{b}\not{c}\not{d}\not{e}).\end{aligned} \tag{A.63}$$

The sign alternates, and the dot product with a marches across the other five vectors. This process can be extended to eight or more γ matrices, but it does get tedious. Traces involving γ^5 are frequently encountered in EWK theory. The trace of γ^5, $\gamma^\mu\gamma^5$, $\gamma^\mu\gamma^\nu\gamma^5$, and $\gamma^\mu\gamma^\nu\gamma^\rho\gamma^5$ all vanish. The first non-zero trace is

$$\mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^5) = 4i\epsilon^{\mu\nu\rho\sigma}, \quad \epsilon^{0123} = -1. \tag{A.64}$$

A property of Dirac algebra called Fierz reordering was discussed relative to the V-A Interaction in Chapter 7.

A.11. Spin Effects

The trace calculations are based on a sum over all polarization states. In the text, polarized initial and final states have been treated in two ways. One, used for example in the annihilation of positronium in Chapter 9, utilizes the explicit form of the spinors without taking traces. The other method, widely used in the text, takes traces as usual, but inserts spin projection operators. For spin 1/2 massive particles, like electrons, the spin direction is defined in the particle rest frame:

$$\begin{aligned}&s^\mu = (0,\hat{n}), \quad s_\mu s^\mu = -1, \quad s^\mu p_\mu = 0,\\ &\hat{n} = \cos\theta\hat{z} + \sin\theta\cos\phi\hat{x} + \sin\theta\sin\phi\hat{y}, \quad \langle\vec{\sigma}\cdot\hat{n}\rangle = +1.\end{aligned} \tag{A.65}$$

The four vector s^μ has no time component in the rest frame, and is a space-like vector. A Lorentz boost along the $\hat{z}$ direction leaves s_x

and s_y invariant, so we focus on longitudinal polarization, $s_z = +1$. Then for the momentum and spin we have:

$$\Lambda^\mu_\nu p^\nu = \begin{pmatrix} \gamma & 0 & 0 & v\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ v\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma m \\ 0 \\ 0 \\ v\gamma m \end{pmatrix}, \quad \text{and}$$

$$\Lambda^\mu_\nu s^\nu = \begin{pmatrix} v\gamma \\ 0 \\ 0 \\ \gamma \end{pmatrix} = \begin{pmatrix} p/m \\ 0 \\ 0 \\ E/m \end{pmatrix}. \tag{A.66}$$

The Lorentz transformed spin four vector explicitly involves the mass, so we may not set $m \to 0$, and it is upside down relative to the momentum four vector. Projection operators for helicity ±1 are

$$\Lambda_\pm = \frac{1 \pm \gamma^5 \not{s}}{2}. \tag{A.67}$$

In order to calculate the probability of positive helicity, simply multiply the completeness operator by Λ_+, thus,

$$\sum_{\text{spins}} u^s(p)\bar{u}^s(p) \to (\not{p} + m)\frac{1 + \gamma^5 \not{s}}{2}. \tag{A.68}$$

Since $\left[\gamma^5 \not{s}, \not{p}\right] = 0$, the order of the operators does not matter. Equation (6.49) gives a useful reduction formula for the $V - A$ weak interaction. This is part of the moun tensor for $\mu \to e\nu\bar{\nu}$:

$$\begin{aligned} &(1 - \gamma^5)(\not{p}_1 + m)\frac{1 + \lambda_1 \gamma^5 \not{s}_1}{2}(1 + \gamma^5) \\ &= (1 - \gamma^5)\frac{\not{p}_1 - \lambda_1 m \not{s}_1 + m}{2}(1 + \gamma^5). \end{aligned} \tag{A.69}$$

Hence we may substitute $p_1 \to (p_1 - \lambda_1 m s_1)/2$, where $\lambda_1 = \pm 1$ for the two helicity states, in the final expression.

For massless Dirac particles, like neutrinos, there is no rest frame, so the definitions that we have proposed do not work. In this case the helicity projection operators are:

$$\Lambda_{\pm} = \frac{1 \pm \gamma^5}{2}, \tag{A.70}$$

For massive vector particles like $W^{\pm}$ and Z, the three-component spin vector is defined in the rest frame in analogy with the $l = 1$ spherical harmonics:

$$\begin{aligned} \epsilon^{\mu,\lambda=\pm 1} &= (0, \mp 1/\sqrt{2}, -i/\sqrt{2}, 0), \\ \epsilon^{\mu,\lambda=0} &= (0, 0, 0, 1), \quad \epsilon^{\lambda}_{\mu} p^{\mu} = 0. \end{aligned} \tag{A.71}$$

A Lorentz boost then leaves $\epsilon^{\mu,\lambda=\pm 1}$ invariant, and $\epsilon^{\mu,\lambda=0} \rightarrow (p/m, 0, 0, E/m)$ just like spin 1/2. The completeness sum is given in Equation (A.44) for the W boson.

Real photons have no rest frame, and have only two spin components, analogous to $\epsilon^{\mu,\lambda=\pm 1}$, with $\epsilon^{\mu,\lambda=0}$ absent. Thanks to current conservation, the effective sum over photon polarization states in QED is

$$\sum_{\lambda} \epsilon^{*\lambda}_{\mu} \epsilon^{\lambda}_{\nu} = -g_{\mu\nu}. \tag{A.72}$$

Virtual photons can have longitudinal polarizations, as discussed in Chapter 5 in electron–proton elastic scattering.

References

[1] P. A. Zyla *et al.* (Particle Data Group), *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020).

[2] V. Barger and R. Phillips, *Collider Physics* (Addison Wesley, New York, 1997).

Index

$\Delta Q = \Delta S$ rule, 30, 49, 460
$\Delta^{++} \to p + \pi^+$, 10
Λ hyperon, 7, 29, 284
Ω^- decay modes, 285
$\Upsilon(1S)$, $(2S)$, $(3S)$, $(4S)$, 436, 438
β rays, 1
η, 8
γ^5, 89, 171
ν_μ, 14
ν_{e}, 14
ϕ, 28
π meson, 2, 3
$\pi^+ \to \pi^0 + e^+ + \nu_e$, 280
π_e/π_μ branching ratio, 280
τ lepton, 338
$\lambda(x, y, z)$, 63

Abbey, Westminster, 73
Abelian gauge theory, 140
action integral, 106
AGS synchrotron, 14
alpha particles, 1
Altarelli–Parisi equations, 232, 249
ambiguities, 483
Anderson, Carl, 2
angular equation, 422
angular momentum algebra, 151
annihilation $e^-e^+ \to \gamma + \gamma$, 427
anomalous magnetic moment, 276
anomalous part of the proton magnetic moment, 218
anticommutation rules, 73
anticommutator, 141, 143
antisymmetric four-dimensional ϵ tensor, 57
associated Laguerre polynomial, 423
associated Legendre function, 15
asymmetry parameter, 276, 363
atomic mass units, 258
axial vector, 90

$B_d = d\bar{b}$ mixing, 464
b tag, 345
BaBar, 478
back scattering of laser photons, 207
Barger and Phillips, 58, 270, 274, 275, 352, 374, 385, 403, 418
baryon current, 219
baryon magnetic moment, 49
beam method for neutron lifetime, 269
Becquerel, Henri, 1, 257
Belle experiment, 301, 466, 478
BESIII, 286

beta decay, 257, 258
  Λ, 290
  tritium, 292
Bethe–Heitler mechanism, 245
Bhabha scattering, 38, 183, 186, 189
Bjorken and Drell, 57, 207
Bjorken, J. D., 229, 230, 231
Bjorken–Drell metric, 58
Bjorken x, 214
Bohr's formula, 82, 422
Born approximation, 131, 488
Bose–Einstein statistics, 115, 199
bottle method for neutron lifetime, 269
bottom mesons, 44
bottom quark, 31, 43
bottomonium, 125, 435
bound states of $Q\bar{Q}$ pairs, 434
box diagrams, 340
box normalization, 76
box normalized wave functions, 130
Boyarski, A. M., 41, 42
Braaten and Levielle, 407
Bragg reflection, 196
Breit–Wigner form, 12
Brookhaven muon $g-2$ experiment, 209
Brookhaven National Laboratory, 14, 42, 435, 451
Butler, Clifford, 6

Cabibbo angle, 43, 259, 261, 269, 292, 315
Cabibbo, Nicola, 232, 294, 339
calcite crystal, 196
canonical momentum, 108
CDF top quark candidate, 341
CDHS neutrino experiment, 232, 320
CERN library, 331
CESR e^-e^+ storage ring, 438
charge asymmetry, 327, 460
charge conjugation, 92
  wave function, 92
charge conjugation operator, 31
charge independence, 4, 5
charge rule, 389
charge symmetry, 4
charged field, 115, 152, 153
charged pions, 3
charged vector boson
  magnetic moment, 417
charm quark, 31
charmed baryon, 298
charmed mesons, 44
charmonium, 125, 438
Chau, Ling Lie, 473, 475
circular polarization states, 123
CKM 'golden' triangle, 472
CKM matrix, 294, 339, 340, 390, 468, 473
Clebsch–Gordan coefficients, 5, 15, 28
CLEO, 299
CMS collaboration, 408, 414
coherent state, 148
coherent states in ket language, 150
Coleman, Sidney, 388
color factor, 39
color SU(3), 388
color symmetry, 29
commutation rules, 123
completeness for wave functions, 132
completeness relation for a spin one massive boson, 325
completeness sums for electron and positron spinors, 86
complex k' plane, 133
Compton absorption in magnetized iron, 206
Compton back scattering, 359
Compton scattering, 178, 196, 198
  Feynman diagram, 199
Compton wave length, 36, 197
conservation laws, 3
conservation of a current, 70
constituent quark magnetic moments, 225
constituent quark masses, 44, 225
'current quark' masses, 225
continuity equation, 69
continuous space time, 125
continuum normalization, 76, 110
contraction, 115

cosmic ray, 6
cosmic ray neutrinos, 310
Coulomb pole, 180
Coulomb potential energy, 130
coupling of the Z boson to fermions, 348
covariant derivative, 139, 382
  $SU(2)_L \times U(1)_Y$, 392
  non-Abelian gauge theory, 384
covariant perturbation theory, 160
Cowan and Reines, 310
CP violation, 340, 454, 462, 464, 468, 477, 479, 480, 483
Cronin and Overseth, 286
cross couples the creation and destruction operators, 116
cross-section, 130, 161
  Schrödinger's quantum mechanics, 124
crossed amplitude, 181
crossing symmetry, 179
CTEQ collaboration, 242, 251
current densities, 394
current quark, 46, 226
cyclotron frequency, 80, 208
cyclotrons, 6

D mesons, 295
Davis, Ray, 310
Dalitz, R. H., 33
decay rate
  top quark, 344
  unstable particle, 163
deep inelastic scattering, 232
degeneracy, 18
delta function, 66, 109
density of final states, 67, 128
deuteron, 5
differential cross-section, 162
dimensional regularization, 209
dimuon signal at the Higgs mass, 408
'dipole' formula, 222
Dirac algebra in d dimensions, 209
Dirac current modification, 218
Dirac equation, 71, 82, 84, 153
  electron spinors, 75, 77
  external electromagnetic field, 77, 78
  Lorentz transformation, 86
  positron spinor, 85
Dirac field, 140
Dirac wave function, 72
Dirac's relativistic equation, 198
Dirac, Paul, 2, 73, 125, 126
'direct' CP violation, 459
discovery of the Higgs, 369
discreet space time, 125
double scattering experiment, 290
Drell–Yan mechanism, 245, 323, 329, 404
dynamical symmetry breaking, 294, 394

$e^- + e^+ \to \mu^- + \mu^+$, 163
$e^-\mu^- \to e^-\mu^-$, 179, 215, 229
$e^-e^+ \to \mu^-\mu^+$, 172
$e^-e^+ \to e^-e^+$, 164
Einstein's velocity addition theorem, 96
Einstein, Albert, 1
elastic electron scattering, 213
electric field, 123
electric form factor, 221
electromagnetic wave, 117
electromagnetism, 3
electron
  field operator, 165
  g factor, 80
  propagator, 144, 183
  spin, 190
electron–antineutrino angular correlation, 264
electron–positron collider, 359
Electron–Positron Storage Rings, 34
electron–proton elastic scattering, 217
electron–proton inelastic scattering, 224
electroweak phenomenology, 309
electroweak radiative corrections, 261

electroweak theory, 259, 260, 294, 311, 339, 346, 348, 381, 394
emission, 2
equal spacing rule, 25
equal time commutator, 143
Euler angles, 53
'exact' energy, 82
exotic Coulomb atoms, 433
exotic mesons, 301

$F^{\mu\nu}$, 119
Fermi coupling constant, 259, 261
Fermi, Enrico, 2, 160, 263, 266
Fermilab, 314, 326, 435
Feynman diagram, 181, 270
Feynman propagator for the scalar field, 135
Feynman, Richard, 81, 230–232, 234
Feynman rules, 115, 168, 181
Feynman's 'slash' notation, 74
Fierz transformation, 374
fine structure constant, 36
fine structure splitting, 83
first-order electroweak theory, 349, 362
flavor changing neutral currents, 43
flavor tag, 340, 466
flip amplitudes, 194
form factors, 218
four momentum, 60
four-dimensional antisymmetric tensor, 172
four-dimensional form, 118
four-dimensional Laplacian, 71
four-dimensional volume element, 60
four-quark molecule, 302
Fourier amplitude, 76
Fourier expansion, 75, 76
Fourier series, 108
fragmentation, 39, 235
Frascati, 34
free field Lagrange density, 119
Friedman, Jerome, 230
fuel cycle in the sun, 265
full width, 8
functional derivatives, 106

g factor anomaly, 209
$g-2$, 81
G parity, 8, 47
GaAs photocathode, 359
gamma function, 443
Gamow, George, 257
Gamow–Teller, 264, 266
Gargamelle heavy liquid bubble chamber, 309
gauge condition, 118, 325
gauge invariant, 78, 139
gauge transformation, 78
Gauss' theorem, 120
Gell-Mann, Murray, 20, 232
Gell-Mann–Okubo mass formula, 26, 44
generate events, 65
Glashow, Sheldon Lee, 294, 381
gluon distribution function, 249
gluon splitting function, 250
gluons, 29, 231, 329
Golden Rule #1, 67
Golden Rule #2, 67, 124, 129
Goldstone boson, 397
Gordon decomposition of the vector current, 91, 217
Gordon, Walter, 70, 91
gradient operator, 59
Grant and Page, 207
Grassmann numbers, 141
gravity, 3
Green's function, 132
  Helmholtz's equation, 131
Gupta and Bleuler, 122
Gupta and Bleuler's argument, 155

^{3}He, 225
hadron tensor, 219
hadronization, 39
hadrons, 26
Halzen and Martin, 57, 75, 220, 228

Hamiltonian, 83, 104, 110, 123, 142
  density, 120
  vacuum expectation value, 112
Hamiltonian formulations of mechanics, 78
Hamiltonian operator, 69
harmonic oscillator, 104, 440
harmonic oscillator operators, 115
'hat' variables, 239
heavier quarks, 31
heavy quark mixing, 31
Heisenberg picture, 125
helicity conservation amplitudes, 194
helicity projection operators, 263
HERA, 213, 242
hermeticity, 331
Higgs boson, 326, 381, 435
Higgs boson mass, 402
Higgs couplings, 402
  electrons, 404
Higgs field Lagrangian, 402
Higgs mechanism, 388, 394, 397
Higgs production via $e^-e^+ \to Z + h$, 413
Higgs scalar, 138
high energy neutrinos, 14
high mass lepton pairs in hadronic collisions, 244
Homestake neutrino detector, 370
hybrid calculation, 429
hydrogen atom, 82, 421
hydrogen fine structure, 425
hydrogen hyperfine splitting, 426
hydrogen radial wave function, 422
hyperbolic functions, 95
  identities, 100
hypercharge, 15, 31, 115, 389
hyperfragment, 33
hyperon beta decay, 24

inelastic scattering, 214
infinite momentum frame, 224
infinite well, 440
interaction Hamiltonian, 165
interaction picture, 126
interference terms in QED cross sections, 185
intermediate vector bosons, 322
invariant kinematic variables, 37
invariant mass of the recoil, 214
invariant three-body phase space, 68
I-spin, 4
  commutation rules, 22
  invariance, 246
  singlet ω, 10

J/ψ, 42, 435
jet, 40, 235
JLab, 213

K factor, 329
K^0 meson, 7
K^0 mixing, 448
$K\bar{K}$ mixing, 474
$K_S \to \pi^+\pi^-$ decay amplitude, 282
Katrin collaboration, 292
Kendall, Henry, 230
KEK, 477
KEKB, 477, 478
Keung and Marciano, 413
kinematic boundary, 64
Klein paradox, 84
Klein, Oskar, 70
Klein–Gordon equation, 70, 107, 108, 117, 152, 198, 325
Klein–Gordon Lagrangian, 108
Klein–Nishina formula, 198, 205
Kobayashi, Makoto, 232, 294, 339
Kronecker δ's, 76

Lagrange density, 105–107
  quantum electrodynamics, 139
Lagrange's equations, 106
Lagrangian formulations of mechanics, 78
Lagrangian with W and B fields, 391
Laguerre polynomial, 423
Lamb shift, 83
Lamb shift splitting, 208
Landau and Lifshitz, 422
Landau, L. D., 428

Larmor frequency, 208
lattice gauge theory, 125
Lawrence, Ernest, 6
Lederman, Leon, 14
Lee, T. D., 13
Lee, Oehme, and Yang, 456
left-handed leptons, 263
'left–right' asymmetry, 290
Legendre's differential equation, 10
length contraction, 61
Lenz vector, 422
LEP, 349, 359
LEPII, 333
LHC, 345, 408
LHCb, 295, 302, 464
lifetime, 8
lifetime of the unstable particle, 163
light quark binding, 42
light quarks, 125
line width Γ_W, 336
linear potential, 440
local $U(1)$ gauge transformation, 140
logarithmic potential, 442
longitudinal component, 236
longitudinal polarization, 237, 285
loop integral, 475
Lorentz boost, 100
Lorentz gauge, 118, 120
Lorentz group, 53
Lorentz transformation, 56, 70, 103
  non-collinear, 81
lowering operator, 21, 104
luminosity, 35

Möller scattering, 181, 189, 195, 370
magnetic field, 124
magnetic form factor, 221
magnetic moment, 208
Mandelstam variables, 37, 173, 180, 186, 188, 198, 241
Manhattan Project, 6
Maskawa, Toshihide, 232, 294, 339
mass difference, 283
massive intermediate vector bosons, 245
massless fields, 393
matrix S for a proper Lorentz transformation, 87
matrix elements for the octet, 23
matter–antimatter puzzle, 454
Maxwell's equations, 35, 103, 116
medical tomography, 427
Merzbacher, E., 287
metric tensor, 58
Mexican hat, 396
microcausality, 134
minimum uncertainty, 148
minimum uncertainty wave packet, 150
missing E_T, 331
mistag rate, 479
mixing angle, 393
molybdenum, 196
Mott cross-section, 217, 229
multiplicity, 234
muon, 3
  decay, 269
  decay rate, 272
  field operator, 165
  Larmor frequency, 433
  lifetime, 276
  pair production, 38
muonium, 433

Nachtmann, Otto, 103, 122, 155, 186, 195, 428
natural cutoff, 131
natural units, 35, 36
Neddermeyer, Seth, 2
negative energies, 83
negative energy spinors, 77
negative magnetic moments, 224
nested integrals, 147
neutral K meson, 30
neutral currents, 347
neutral pion, 3, 8
neutrino, 2, 3, 347
  astronomy, 310
  electron scattering, 311
  flavor mixing, 292, 321
  projection operators, 195
  quark scattering, 315

  scattering, 232, 309
neutrons, 1
  lifetime, 268
  magnetic moment, 218
  structure, 213
Newton, Isaac, 73
non-Abelian gauge transformation, 383, 386
non-leptonic decay, 283
non-relativistic quantum mechanics, 125
normal ordering, 113, 124
normalization, 432
nuclear force, 2
nucleons, 2

odd intrinsic parity, 32
Olsen, S. L., 302
'on mass shell' W boson, 323
one unit of R, 39
one-dimensional harmonic oscillators, 151
optical theorem, 11, 452
orbital angular momentum, 55
parity, 32, 87
parity of the π^0, 32
parity violating, 263
Particle Data Group, 45, 285, 339, 348, 398, 417
particle rest frame, 61, 63
parton cross-section, 328
parton distribution functions, 242
parton model, 214, 224, 226, 228, 231, 232, 234, 238, 241, 246, 316, 318, 323
Paul, Stephan, 268
Pauli exclusion principle, 4, 83, 141
Pauli spin matrices, 20
Pauli, Wolfgang, 2
PEPII, 477, 478
perturbation theory, 82, 125
perturbative QCD, 330
Peskin and Schroeder, 58, 103, 194, 207, 209, 250, 386, 418, 432
phase change on the Dirac wave function, 139
phase shift description of elastic scattering, 10
phase space, 67
phase space factor, 130
photon, 1, 29
photon propagator, 134, 138, 180, 195
pion, 3
  beta decay, 261
  decay, 277
  decay rate, 280
  proton scattering, 5
pion–nucleon, 5
Planck's constant, 8, 131
Planck, Max, 1
Poisson distribution, 151
polarization vector for the W boson, 333
polarized cross-section, 195
polarized muons, 272
polarized Møller scattering, 369
Pontecorvo, Bruno, 371
poor duty factor, 231
positive helicity, 206
positron, 2, 83, 192
positron emission, 258
positron spinors, 85
positronium, 426
  annihilation, 431
  decay rate, 433
  fine structure, 427
  singlet state, 428
Powell, Cecil, 3
products of exponentials, 55
propagator, 135
propagator in momentum space, 138
proper time, 61
proton polarization, 285
proton structure, 213
proton–antiproton colliding beams, 310
protons, 1
pseudorapidity, 97
pseudoscalar, 90, 172
PYTHIA, 251

QCD, 388
quantization of the Dirac field, 138
quantization of the electromagnetic field, 121
quantum electrodynamics, 39, 78, 83, 103, 134
quantum entanglement, 103, 114, 429, 467
quantum field theory, 103
quantum mechanical mixing matrix, 449
quantum mechanical state vectors in field theory, 113
quantum theory of angular momentum, 4
quark binding, 317
quark diagram, 282
quark distribution function, 240
quark flow diagrams, 299
quark level cross-section, 327
quark masses, 44, 224
quark mixing, 338, 339
quark model, 17, 26, 29, 39, 298, 311
quarks, 15, 231
quasi-elastic scattering of electrons from deuterium, 222

$R(\alpha, \beta, \gamma)$, 55
radial Schrödinger equation, 421
radiation gauge, 118
Radiation Laboratory at MIT, 6
raising operator, 21, 104
rapidity, 95, 96
rapidity gaps, 96
regeneration, 34, 451, 462
regenerator, 452
Reines and Cowan, 314
relativistic Breit–Wigner, 330
relativistic kinematics, 62
relativistic quantum mechanics, 125
relativistic time dilation, 61, 276
resonances, 8
right circularly polarized light, 206
right-handed antileptons, 263
Rochester, George, 6
Rosenbluth formula, 220, 221, 230
rotations for the spherical harmonics, 55
Rutherford scattering, 129, 180
Rutherford, Ernest, 1, 129

s channel, 38, 164
S matrix, 159, 166
Sakurai, J. J., 58, 81, 84
Salam, Abdus, 381
'scaling', 231
Schiff, L. I., 113, 149
Schrödinger equation, 36, 77, 125
Schrödinger picture, 125
Schrödinger's wave mechanics, 198
Schrödinger, Erwin, 69, 70
Schwartz, Mel, 14
Schwinger, Julian, 151, 232, 447
single top production, 345
SLAC, 40, 213, 230, 231, 435, 477
SLAC Linear Collider, 207, 359
SN1987A, 310
solar neutrinos, 370
space inversion, 87
SPEAR $e^- e^+$ storage ring at SLAC, 35, 435
special relativity, 53
specific forms of the spinors, 175
spectator model of *D* or *B* meson decay, 295
spherical harmonics, 15, 423
spin 3/2 baryons, 8
spin dependent mass term, 45
spin dependent scattering, 287
spin effects, 190
spin four vector, 324
spin orbit force, 81
spin precession frequency, 80
spin projection operator, 89, 192, 274
spinor, 74, 88
splitting functions, 249
spontaneous symmetry breaking, 107, 326, 346, 394
stable odd–odd nucleus, 266
Standard Model, 388

Steinberger, Jack, 14
step function, 137
storage ring, 34
strange particle reactions, 47
strange particles, 6
strangeness, 7, 15, 115, 281
strong coupling constant, 249
strong force, 3
structure functions, 229, 231, 237
SU(2) × U(1) Lagrangian, 388
SU(3), 15, 20
Sudbury Neutrino Observatory, 370
Super-Kamiokande-IV, 370
symmetries, 3
symmetry breaking, 115
symmetry breaking parameter, 283
synchrotron, 7, 34

t channel, 38, 164
't Hooft and Veltman, 209
T matrix, 159, 160, 168
T matrix element, 259, 267
  positronium, 429
Taylor, Richard, 230
TCP theorem, 92, 449
Tevatron, 326, 327, 341
Thomas precession, 81, 427
Thomson, J. J., 1
three-body decays, 64, 98
three-dimensional rotations, 53
three-dimensional symmetric harmonic oscillator, 423
time dilation, 61
time ordered, 134, 136
time ordered product, 137, 144
time proportional transitions, 128
time reversal, 87, 284
Tomonaga, Shinichiro, 232
top production cross-section, 245
top quark, 43, 46, 293, 341, 435
toponium, 344
totally antisymmetric second rank tensor, 119
Touschek, Bruno, 34
trace operation, 171
trace theorems, 170
transition rate, 160, 161
transverse mass, 96, 331
transverse nature of the electromagnetic wave, 117
transverse polarizations, 237
triangular relation, 50
triplet positronium decays into three γ rays, 429
two-body decay kinematics, 97
two-body fixed target, 62
two-body-to-two-body highly relativistic reaction in the center of mass, 177
two-neutrino invariant phase space, 275

U-spin, 21
  commutation rules, 22
UA1, 465
underlying event, 329
unit of R, 179
unitary symmetry, 14, 19, 30
University of Chicago, 7
Upsilon (4S), 298, 436, 438

V-spin, 21
  commutation rules, 22
vacuum expectation value, 396
  products of field operators, 114
vacuum state, 112
valence up quarks in the proton, 251
vector boson vector field, 389
vector meson, 8
vector potential, 117
Veltman, Martinus, 59
virtual photon, 235, 238

W production in $p\bar{p}$ collisions, 326 ff
W polarization in top quark decay, 346
W rapidity, 335
$W^{\pm}$ magnetic moments, 82, 415
Walker, R. C., 223

Ward's identity, 195
wave function collapse, 114
weak I-spin, 292
weak I-spin model, 388
weak decays of strange particles, 281
weak force, 3
weak isospin triplet of W bosons, 348
weak neutral currents, 309
Weinberg angle, 260, 294
Weinberg, Steven, 260, 309, 381, 388
Wick's theorem, 115, 151, 154
WKB approximation, 440, 442
Wolfenstein, L., 340

$X(3872)$, 301

Yang–Mills, 386, 388
Yukawa interaction term, 403
Yukawa meson, 3
Yukawa particles, 6
Yukawa, Hideki, 2

Z decay modes, 366
Z production in e^-e^+ collisions, 356
Z production in $p\bar{p}$ collisions, 364
Z production with polarized electrons, 360
Z resonance, 357
Zitterbewegung, 84
Zweig's rule, 300
Zweig, George, 232, 300